Coupling of Brownian motions and Perelman's $\mathcal{L}$-functional

Kazumasa Kuwada*† and Robert Philipowski‡

Abstract

We show that on a manifold whose Riemannian metric evolves under backwards Ricci flow two Brownian motions can be coupled in such a way that the expectation of their normalized $\mathcal{L}$-distance is non-increasing. As an immediate corollary we obtain a new proof of a recent result of Topping (J. reine angew. Math. 636 (2009), 93–122), namely that the normalized $\mathcal{L}$-transportation cost between two solutions of the heat equation is non-increasing as well.

Keywords: Ricci flow, $\mathcal{L}$-functional, Brownian motion, coupling.
AMS subject classification: 53C44, 58J65, 60J65.

1 Introduction

Let M be a d-dimensional differentiable manifold, $0 \le \bar{\tau}_1 < \bar{\tau}_2 < T$ and $(g(\tau))_{\tau \in [\bar{\tau}_1, T]}$ a complete backwards Ricci flow on M, i.e. a smooth family of Riemannian metrics satisfying

$$\frac{\partial g}{\partial \tau} = 2 \operatorname{Ric}_{g(\tau)} \tag{1}$$

and such that $(M, g(\tau))$ is complete for all $\tau \in [\bar{\tau}_1, T]$. In this situation Perelman [18, Section 7.1] (see also [5, Definition 7.5]) defined the $\mathcal{L}$-functional of a smooth curve $\gamma : [\tau_1, \tau_2] \to M$ (where $\bar{\tau}_1 \le \tau_1 < \tau_2 \le T$) by

$$\mathcal{L}(\gamma) := \int_{\tau_1}^{\tau_2} \sqrt{\tau} \left[|\dot{\gamma}(\tau)|^2_{g(\tau)} + R_{g(\tau)}(\gamma(\tau)) \right] d\tau,$$

where $R_{g(\tau)}(x)$ is the scalar curvature at x with respect to the metric $g(\tau)$.

Denoting by $L(x, \tau_1; y, \tau_2)$ the infimum of $\mathcal{L}(\gamma)$ over smooth curves $\gamma : [\tau_1, \tau_2] \to M$ satisfying $\gamma(\tau_1) = x$ and $\gamma(\tau_2) = y$, and by

$$W^{\mathcal{L}}(\mu, \tau_1; \nu, \tau_2) := \inf_{\pi} \int_{M \times M} L(x, \tau_1; y, \tau_2) \pi(dx, dy)$$

(the infimum is over all probability measures π on $M \times M$ whose marginals are μ and ν) the associated transportation cost between two probability measures μ and ν on M, Topping [23] (see also Lott [14]) obtained the following result:

*Graduate School of Humanities and Sciences, Ochanomizu University, Tokyo 112-8610, Japan. E-mail: kuwada.kazumasa@ocha.ac.jp / Present address: Institut für Angewandte Mathematik, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany. E-mail: kuwada@iam.uni-bonn.de
†Partially supported by the JSPS fellowship for research abroad.
‡Institut für Angewandte Mathematik, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany. E-mail: philipowski@iam.uni-bonn.de

Theorem 1 (Theorem 1.1 in [23]). *Assume that M is compact and that $\bar{\tau}_1 > 0$. Let $u : [\bar{\tau}_1, T] \times M \to \mathbb{R}_+$ and $v : [\bar{\tau}_2, T] \times M \to \mathbb{R}_+$ be two non-negative unit-mass solutions of the heat equation*

$$\frac{\partial u}{\partial \tau} = \Delta_{g(\tau)} u - Ru,$$

where the term Ru comes from the change in time of the volume element. Then the normalized $\mathcal{L}$-transportation cost

$$\bar{\Theta}(t) := 2(\sqrt{\bar{\tau}_2 t} - \sqrt{\bar{\tau}_1 t}) W^{\mathcal{L}}(u(\bar{\tau}_1 t, \cdot) \operatorname{vol}_{g(\bar{\tau}_1 t)}, \bar{\tau}_1 t; v(\bar{\tau}_2 t, \cdot) \operatorname{vol}_{g(\bar{\tau}_2 t)}, \bar{\tau}_2 t) - 2d \left(\sqrt{\bar{\tau}_2 t} - \sqrt{\bar{\tau}_1 t}\right)^2$$

between the two solutions evaluated at times $\bar{\tau}_1 t$ resp. $\bar{\tau}_2 t$ is a non-increasing function of $t \in [1, T/\bar{\tau}_2]$.

By $g(\tau)$-Brownian motion, we mean the time-inhomogeneous diffusion process whose generator is $\Delta_{g(\tau)}$. As in the time-homogeneous case, the heat distribution $u(\tau, \cdot) \operatorname{vol}_{g(\tau)}$ is expressed as the law of a $g(\tau)$-Brownian motion at time τ. In view of this strong relation between heat equation and Brownian motion, it is natural to ask whether one can couple two Brownian motions on M in such a way that a pathwise analogue of this result involving the function

$$\Theta(t, x, y) := 2\left(\sqrt{\bar{\tau}_2 t} - \sqrt{\bar{\tau}_1 t}\right) L(x, \bar{\tau}_1 t; y, \bar{\tau}_2 t) - 2d\left(\sqrt{\bar{\tau}_2 t} - \sqrt{\bar{\tau}_1 t}\right)^2.$$

holds. The main result of this paper answers it affirmatively as follows:

Theorem 2. *Assume that M has bounded curvature tensor, i.e.*

$$\sup_{x \in M, \tau \in [\bar{\tau}_1, T]} |\operatorname{Rm}_{g(\tau)}|_{g(\tau)}(x) < \infty. \tag{2}$$

Then given any points $x, y \in M$ and any $s \in [1, T/\bar{\tau}_2]$, there exist two coupled $g(\tau)$-Brownian motions $(X_\tau)_{\tau \in [\bar{\tau}_1 s, T]}$ and $(Y_\tau)_{\tau \in [\bar{\tau}_2 s, T]}$ with initial values $X_{\bar{\tau}_1 s} = x$ and $Y_{\bar{\tau}_2 s} = y$ such that the process $(\Theta(t, X_{\bar{\tau}_1 t}, Y_{\bar{\tau}_2 t}))_{t \in [s, T/\bar{\tau}_2]}$ is a supermartingale. In particular $\mathbb{E}[\Theta(t, X_{\bar{\tau}_1 t}, Y_{\bar{\tau}_2 t})]$ is non-increasing. In addition, we can take them so that the map $(x, y) \mapsto (X, Y)$ is measurable.

Remark 1. Obviously, (2) is satisfied if M is compact. Thus it includes the case of Theorem 1. In addition, there are plenty of examples of backwards Ricci flow satisfying (2) even when M is non-compact. Indeed, given a metric g_0 on M with bounded curvature tensor, there exists a unique solution to the Ricci flow $\partial_t g(t) = -2 \operatorname{Ric}_{g(t)}$ with initial condition g_0 satisfying (2) for a short time (see [21] for existence and [4] for uniqueness). Then the corresponding backwards Ricci flow is obtained by time-reversal.

Remark 2. As shown in [12], under backwards Ricci flow $g(\tau)$-Brownian motion cannot explode. Hence $\Theta(t, X_t, Y_t)$ is well-defined for all $t \in [s, T/\bar{\tau}_2]$. This fact also ensures that $u(\tau, \cdot) \operatorname{vol}_{g(\tau)}$ has unit mass whenever it does at the initial time.

Using Theorem 2 we can prove Topping's result even in the non-compact case.

Theorem 3. *Assume that (2) holds. Then the same assertion as in Theorem 1 holds true for nonegative unit mass solutions u and v to the heat equation and the associated functional $\bar{\Theta}(t)$.*

Proof of Theorem 3 using Theorem 2. Fix $1 \le s < t \le T/\bar{\tau}_2$, and let π be an optimal coupling of $u(\bar{\tau}_1 s, \cdot) \operatorname{vol}_{g(\bar{\tau}_1 s)}$ and $v(\bar{\tau}_2 s, \cdot) \operatorname{vol}_{g(\bar{\tau}_2 s)}$. (Existence of an optimal coupling follows from [24, Theorem 4.1], using the obvious lower bound $L(x, \tau_1; y, \tau_2) \ge \frac{2}{3}(\tau_2^{3/2} - \tau_1^{3/2}) \inf_{x \in M, \tau \in [\bar{\tau}_1, T]} R_{g(\tau)}(x)$.) For each $(x, y) \in M \times M$, we take coupled Brownian motions $(X^x_\tau)_{\tau \in [\bar{\tau}_1 s, T]}$ and $(Y^y_\tau)_{\tau \in [\bar{\tau}_2 s, T]}$ with initial values $X^x_{\bar{\tau}_1 s} = x$ and $Y^y_{\bar{\tau}_2 s} = y$ as in Theorem 2. Since $(x, y) \mapsto (X^x, Y^y)$ is measurable, we can construct a coupling of two Brownian motions (X, Y) with initial distribution π by following a usual manner. Then the joint distribution of $X_{\bar{\tau}_1 t}$ and $Y_{\bar{\tau}_2 t}$ is a coupling of $u(\bar{\tau}_1 t, \cdot) \operatorname{vol}_{g(\bar{\tau}_1 t)}$ and $v(\bar{\tau}_2 t, \cdot) \operatorname{vol}_{g(\bar{\tau}_2 t)}$, so that $\bar{\Theta}(t) \le \mathbb{E}[\Theta(t, X_{\bar{\tau}_1 t}, Y_{\bar{\tau}_2 t})] \le \mathbb{E}[\Theta(s, X_{\bar{\tau}_1 s}, Y_{\bar{\tau}_2 s})] = \bar{\Theta}(s)$. $\square$

2 Remarks concerning related work

The Ricci flow was introduced by Hamilton [8]. There he effectively used it to solve the Poincaré conjecture for 3-manifolds with positive Ricci curvature. By following his approach, Perelman [18, 19, 20] finally solved the Poincaré conjecture (see also [3, 9, 17]). There he used $\mathcal{L}$-functional as a crucial tool. At the same stage, he also studied the heat equation in [18] in relation with the geometry of Ricci flows. It suggests that analysing the heat equation is still an efficient way to investigate geometry of the underlying space even in the time-dependent metric case. This general principle has been confirmed in recent developments in this direction. In connection with the theory of optimal transportation, McCann and Topping [16] showed contraction in the L^2-Wasserstein distance for the heat equation under backwards Ricci flow on a compact manifold. Topping's result [23] can be regarded as an extension of it to contraction in the normalized $\mathcal{L}$-transportation cost (see [14] also). By taking $\bar{\tau}_2 \to \bar{\tau}_1$, he recovered the monotonicity of Perelman's $\mathcal{W}$-entropy, which is one of fundamental ingredients in Perelman's work.

A probabilistic approach to these problems is initiated by Arnaudon, Coulibaly and Thalmaier. In [2, Section 4], they sharpened McCann and Topping's result [16] to a pathwise contraction in the following sense: There is a coupling $(X_t, Y_t)_{t\geq 0}$ of two Brownian motions starting from $x, y \in X$ respectively such that the $g(t)$-distance between X_t and Y_t is non-increasing in t almost surely. In their approach, probabilistic techniques based on analysis of sample paths made it possible to establish such a pathwise estimate. It should be mentioned that, as another advantage of their approach, their argument works even on non-compact M (cf. [12]). Our approach is the same as theirs in spirit. In fact, such advantages are also inherited to our results. Unfortunately, we cannot expect a pathwise contraction as theirs since our problem differs in nature from what is studied in [1] (see Remark 7). However, it should be noted that this new fact is revealed as a result of our pathwise arguments. Furthermore, we can expect that our approach makes it possible to employ several techniques in stochastic analysis to obtain more detailed behavior of $\Theta(t, X_{\bar{\tau}_1 t}, Y_{\bar{\tau}_2 t})$, especially in the limit $\bar{\tau}_2 \to \bar{\tau}_1$, in a future development. Note that, from technical point of view, our method relies on the result in [11] and it is different from Arnaudon, Coulibaly and Thalmaier's one.

3 Coupling of Brownian motions in the absence of $\mathcal{L}$-cut locus

Since the proof of Theorem 2 involves some technical arguments, first we study the problem in the case that the $\mathcal{L}$-distance L has no singularity. More precisely,

Assumption 1. The $\mathcal{L}$-cut locus is empty.

See subsection 5.1 or [5, 23, 25] for the definition of $\mathcal{L}$-cut locus. Under Assumption 1, the following holds:

1. For all $x, y \in M$ and all $\bar{\tau}_1 \leq \tau_1 < \tau_2 \leq T$ there is a unique minimizer $\gamma_{xy}^{\tau_1\tau_2}$ of $L(x, \tau_1; y, \tau_2)$ (existence of $\gamma_{xy}^{\tau_1\tau_2}$ is proved in [5, Lemma 7.27], while uniqueness follows immediately from the characterization of $\mathcal{L}$-cut locus, see subsection 5.1).
2. The function L is globally smooth.

Thus, in this case, we can freely use stochastic analysis on the frame bundle without taking any care on regularity of L. In section 5, we present the complete proof of Theorem 2 using a random walk approximation (see Remark 8 for further details on the choice of our approach).

3.1 Construction of the coupling

A $g(\tau)$-Brownian motion $\tilde{X}$ on M (scaled in time by the factor $\bar{\tau}_1$) starting at a point $x \in M$ at time $s \in [1, T/\bar{\tau}_2]$ can be constructed in the following way [1, 7, 12]: Let $\pi : \mathcal{F}(M) \to M$ be the frame bundle and $(e_i)_{i=1}^d$ the standard basis of $\mathbb{R}^d$. For each $\tau \in [\bar{\tau}_1, T]$ let $(H_i(\tau))_{i=1}^d$ be the associated $g(\tau)$-horizontal vector fields on $\mathcal{F}(M)$ (i.e. $H_i(\tau, u)$ is the $g(\tau)$-horizontal lift of ue_i). Moreover let $(\mathcal{V}^{\alpha,\beta})_{\alpha,\beta=1}^d$ be the canonical vertical vector fields, i.e. $(\mathcal{V}^{\alpha,\beta} f)(u) := \frac{\partial}{\partial m_{\alpha\beta}}\Big|_{\mathbf{m}=\mathrm{Id}} (f(u(\mathbf{m})))$ ($\mathbf{m} = (m_{\alpha\beta})_{\alpha,\beta=1}^d \in GL_d(\mathbb{R})$), and let $(W_t)_{t\geq 0}$ be a standard $\mathbb{R}^d$-valued Brownian motion. By $\mathcal{O}^{g(\tau)}(M)$, we denote the $g(\tau)$-orthonormal frame bundle.

We first define a horizontal Brownian motion on $\mathcal{F}(M)$ as the solution $\tilde{U} = (\tilde{U}_t)_{t\in[s,T/\bar{\tau}_1]}$ of the Stratonovich SDE

$$d\tilde{U}_t = \sqrt{2\bar{\tau}_1} \sum_{i=1}^d H_i(\bar{\tau}_1 t, \tilde{U}_t) \circ dW_t^i - \bar{\tau}_1 \sum_{\alpha,\beta=1}^d \frac{\partial g}{\partial \tau}(\bar{\tau}_1 t)(\tilde{U}_t e_\alpha, \tilde{U}_t e_\beta) \mathcal{V}^{\alpha\beta}(\tilde{U}_t) dt \tag{3}$$

with initial value $\tilde{U}_s = u \in \mathcal{O}_x^{g(\bar{\tau}_1 s)}(M)$, and then define a scaled Brownian motion $\tilde{X}$ on M as

$$\tilde{X}_t := \pi \tilde{U}_t.$$

Note that $\tilde{X}_t$ does not move when $\bar{\tau}_1 = 0$. The last term in (3) ensures that $\tilde{U}_t \in \mathcal{O}^{g(\bar{\tau}_1 t)}(M)$ for all $t \in [s, T/\bar{\tau}_1]$ (see [1, Proposition 1.1], [7, Proposition 1.2]), so that by Itô's formula for all smooth $f : [s, T/\bar{\tau}_1] \times M \to \mathbb{R}$

$$df(t, \tilde{X}_t) = \frac{\partial f}{\partial t}(t, \tilde{X}_t) dt + \sqrt{2\bar{\tau}_1} \sum_{i=1}^d (\tilde{U}_t e_i) f(t, \tilde{X}_t) dW_t^i + \bar{\tau}_1 \Delta_{g(\tau_1 t)} f(t, \tilde{X}_t) dt.$$

Let us define $(X_\tau)_{\tau\in[\bar{\tau}_1 s, T]}$ by $X_{\bar{\tau}_1 t} := \tilde{X}_t$. Then X_τ becomes a $g(\tau)$-Brownian motion when $\bar{\tau}_1 > 0$.

Remark 3. Intuitively, it might be helpful to think that X_τ lives in $(M, g(\tau))$, or $\tilde{X}_t$ lives in $(M, g(\bar{\tau}_1 t))$. The same is true for Y and $\tilde{Y}$ which will be defined below. Similarly, for all curves $\gamma : [\tau_1, \tau_2] \to M$ in this paper, we can naturally regard $\gamma(\tau)$ as in $(M, g(\tau))$.

We now want to construct a second scaled Brownian motion $\tilde{Y}$ on M in such a way that its infinitesimal increments $d\tilde{Y}_t$ are "space-time parallel" to those of $\tilde{X}$ along the minimal $\mathcal{L}$-geodesic (namely, the minimizer of L) from $(\tilde{X}_t, \bar{\tau}_1 t)$ to $(\tilde{Y}_t, \bar{\tau}_2 t)$. To make this idea precise, we first define the notion of space-time parallel vector field:

Definition 1 (space-time parallel vector field)**.** Let $\bar{\tau}_1 \leq \tau_1 < \tau_2 \leq T$ and $\gamma : [\tau_1, \tau_2] \to M$ be a smooth curve. We say that a vector field Z along γ is *space-time parallel* if

$$\nabla_{\dot{\gamma}(\tau)}^{g(\tau)} Z(\tau) = -\operatorname{Ric}_{g(\tau)}^{\#}(Z(\tau)) \tag{4}$$

holds for all $\tau \in [\tau_1, \tau_2]$. Here $\nabla^{g(\tau)}$ stands for the covariant derivative associated with the $g(\tau)$-Levi-Civita connection and $\operatorname{Ric}_{g(\tau)}^{\#}$ is defined by regarding the $g(\tau)$-Ricci curvature as a (1,1)-tensor.

Remark 4. Since (4) is a linear first-order ODE, for any $\xi \in T_{\gamma(\tau_1)}M$ there exists a unique space-time parallel vector field Z along γ with $Z(\tau_1) = \xi$.

Remark 5. Whenever Z and Z' are space-time parallel vector fields along a curve γ, their $g(\tau)$-inner product is constant in τ:

$$\begin{aligned}\frac{d}{d\tau}\langle Z(\tau), Z'(\tau)\rangle_{g(\tau)} &= \frac{\partial g}{\partial \tau}(\tau)(Z(\tau), Z'(\tau)) + \langle \nabla^{g(\tau)}_{\dot\gamma(\tau)} Z(\tau), Z'(\tau)\rangle_{g(\tau)} + \langle Z(\tau), \nabla^{g(\tau)}_{\dot\gamma(\tau)} Z'(\tau)\rangle_{g(\tau)} \\ &= 2\,\mathrm{Ric}_{g(\tau)}(Z(\tau), Z'(\tau)) - \mathrm{Ric}_{g(\tau)}(Z(\tau), Z'(\tau)) - \mathrm{Ric}_{g(\tau)}(Z(\tau), Z'(\tau)) \\ &= 0.\end{aligned}$$

Remark 6. The emergence of the Ricci curvature in (4) is based on the Ricci flow equation (1). Indeed, we can generalize the notion of space-time parallel transport even in the absence of (1) with keeping the property in the last remark. This would be a natural extension in the sense that it coincides with the usual parallel transport when $g(\tau)$ is constant in τ. On the other hand, it is convenient to define it as (4) for later use in this paper.

Definition 2 (space-time parallel transport)**.** For $x, y \in M$ and $\bar\tau_1 \le \tau_1 < \tau_2 \le T$, we define a map $m^{\tau_1\tau_2}_{xy} : T_xM \to T_yM$ as follows: $m^{\tau_1\tau_2}_{xy}(\xi) := Z(\tau_2)$, where Z is the unique space-time parallel vector field along $\gamma^{\tau_1\tau_2}_{xy}$ with $Z(\tau_1) = \xi$. By Remark 5, $m^{\tau_1\tau_2}_{xy}$ is an isometry from $(T_xM, g(\tau_1))$ to $(T_yM, g(\tau_2))$. In addition, it smoothly depends on x, τ_1, y, τ_2 under Assumption 1.

We now define a second horizontal scaled Brownian motion $\tilde V = (\tilde V_t)_{t\in[s,T/\bar\tau_2]}$ on $\mathcal{F}(M)$ as the solution of

$$d\tilde V_t = \sqrt{2\bar\tau_2}\sum_{i=1}^d H^*_i(\tilde U_t, \bar\tau_1 t; \tilde V_t, \bar\tau_2 t) \circ dW^i_t - \bar\tau_2 \sum_{\alpha,\beta=1}^d \frac{\partial g}{\partial\tau}(\bar\tau_2 t)(\tilde V_t e_\alpha, \tilde V_t e_\beta)\mathcal{V}^{\alpha\beta}(\tilde V_t)dt$$

with initial value $\tilde V_s = v \in \mathcal{O}^{g(\bar\tau_2 s)}_y(M)$, and we set $\tilde Y_t := \pi\tilde V_t$. Here $H^*_i(u, \tau_1; v, \tau_2)$ is the $g(\tau_2)$-horizontal lift of $ve^*_i(u, \tau_1; v, \tau_2)$, where

$$e^*_i(u, \tau_1; v, \tau_2) := v^{-1} m^{\tau_1,\tau_2}_{\pi u, \pi v} u e_i.$$

As we did for $\tilde X$, let us define $(Y_\tau)_{\tau\in[\bar\tau_2 s, T]}$ by $Y_{\bar\tau_2 t} := \tilde Y_t$ to make Y a $g(\tau)$-Brownian motion. From theoretical point of view, it seems to be natural to work with (X_τ, Y_τ) (see Remark 3). However, for technical simplicity, we will prefer to work with $(\tilde X_t, \tilde Y_t)$ instead in the sequel.

3.2 Proof of Theorem 2 in the absence of $\mathcal{L}$-cut locus

Our argument in this section is based on the following Itô's formula for $(\tilde X_t, \tilde Y_t)$.

Lemma 1. *Let f be a smooth function on $[s, T/\bar\tau_2] \times M \times M$. Then*

$$\begin{aligned}df(t, \tilde X_t, \tilde Y_t) &= \frac{\partial f}{\partial t}(t, \tilde X_t, \tilde Y_t)dt + \sum_{i=1}^d \left[\sqrt{2\bar\tau_1}\tilde U_t e_i \oplus \sqrt{2\bar\tau_2}\tilde V_t e^*_i\right] f(t, \tilde X_t, \tilde Y_t) dW^i_t \\ &\quad + \sum_{i=1}^d \mathrm{Hess}_{g(\bar\tau_1 t)\oplus g(\bar\tau_2 t)} f\big|_{(t,\tilde X_t,\tilde Y_t)} \left(\sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e^*_i, \sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e^*_i\right) dt.\end{aligned}$$

Here the Hessian of f is taken with respect to the product metric $g(\bar\tau_1 t) \oplus g(\bar\tau_2 t)$, e^*_i stands for $e^*_i(\tilde U_t, \bar\tau_1 t; \tilde V_t, \bar\tau_2 t)$, and for tangent vectors $\xi_1 \in T_xM$, $\xi_2 \in T_yM$ we write $\xi_1 \oplus \xi_2 := (\xi_1, \xi_2) \in T_{(x,y)}(M \times M)$.

Proof. Itô's formula applied to a smooth function $\tilde f$ on $[s, T/\bar\tau_2] \times \mathcal{F}(M) \times \mathcal{F}(M)$ gives

$$\begin{aligned}
&d\tilde f(t, \tilde U_t, \tilde V_t) \\
&= \frac{\partial \tilde f}{\partial t}(t, \tilde U_t, \tilde V_t)dt + \sum_{i=1}^{d} \left[\sqrt{2\bar\tau_1} H_i(\bar\tau_1 t, \tilde U_t) \oplus \sqrt{2\bar\tau_2} H_i^*(\tilde U_t, \bar\tau_1 t; \tilde V_t, \bar\tau_2 t)\right] \tilde f(t, \tilde U_t, \tilde V_t) dW_t^i \\
&\quad + \sum_{i=1}^{d} \left[\sqrt{\bar\tau_1} H_i(\bar\tau_1 t, \tilde U_t) \oplus \sqrt{\bar\tau_2} H_i^*(\tilde U_t, \bar\tau_1 t; \tilde V_t, \bar\tau_2 t)\right]^2 \tilde f(t, \tilde U_t, \tilde V_t) dt \\
&\quad - \sum_{\alpha,\beta=1}^{d} \left[\bar\tau_1 \frac{\partial g}{\partial \tau}(\bar\tau_1 t)(\tilde U_t e_\alpha, \tilde U_t e_\beta) \mathcal{V}^{\alpha\beta}(\tilde U_t) \oplus \bar\tau_2 \frac{\partial g}{\partial \tau}(\bar\tau_2 t)(\tilde V_t e_\alpha, \tilde V_t e_\beta) \mathcal{V}^{\alpha\beta}(\tilde V_t)\right] \tilde f(t, \tilde U_t, \tilde V_t) dt.
\end{aligned}$$

The claim follows by choosing $\tilde f(t, u, v) := f(t, \pi u, \pi v)$ because the function considered here is constant in the vertical direction so that the term involving $\mathcal{V}^{\alpha\beta} \tilde f$ vanishes. □

Let $\Lambda(t, x, y) := L(x, \bar\tau_1 t; y, \bar\tau_2 t)$. In order to apply Lemma 1 to the function Θ we need the following proposition whose proof is given in the next section:

Proposition 1. *Take $x, y \in M$, $u \in \mathcal{O}_x^{g(\bar\tau_1 t)}(M)$ and $v \in \mathcal{O}_y^{g(\bar\tau_2 t)}(M)$. Let γ be a minimizer of $L(x, \bar\tau_1 t; y, \bar\tau_2 t)$. Assume that $(x, \bar\tau_1 t; y, \bar\tau_2 t)$ is not in the $\mathcal{L}$-cut locus. Set $\xi_i := \sqrt{\bar\tau_1} u e_i \oplus \sqrt{\bar\tau_2} v e_i^*(u, \bar\tau_1 t; v, \bar\tau_2 t)$. Then*

$$\begin{aligned}
\frac{\partial \Lambda}{\partial t}(t, x, y) = \frac{1}{t} \int_{\bar\tau_1 t}^{\bar\tau_2 t} \tau^{3/2} \bigg(&\frac{3}{2\tau} R_{g(\tau)}(\gamma(\tau)) - \Delta_{g(\tau)} R_{g(\tau)}(\gamma(\tau)) - 2 \,|\operatorname{Ric}_{g(\tau)}|_{g(\tau)}^2(\gamma(\tau)) \\
&- \frac{1}{2\tau} |\dot\gamma(\tau)|_{g(\tau)}^2 + 2 \operatorname{Ric}_{g(\tau)}(\dot\gamma(\tau), \dot\gamma(\tau)) \bigg) d\tau, \qquad (5)
\end{aligned}$$

$$\begin{aligned}
&\sum_{i=1}^{d} \operatorname{Hess}_{g(\tau_1) \oplus g(\tau_2)} \Lambda \Big|_{(t,x,y)} (\xi_i, \xi_i) \\
&\quad \le \frac{d\sqrt{\tau}}{t} \bigg|_{\tau=\bar\tau_1 t}^{\tau=\bar\tau_2 t} + \frac{1}{t} \int_{\bar\tau_1 t}^{\bar\tau_2 t} \tau^{3/2} \Big(2 \left|\operatorname{Ric}_{g(\tau)}\right|_{g(\tau)}^2 (\gamma(\tau)) + \Delta_{g(\tau)} R_{g(\tau)}(\gamma(\tau)) \\
&\qquad - \frac{2}{\tau} R_{g(\tau)}(\gamma(\tau)) - 2 \operatorname{Ric}_{g(\tau)}(\dot\gamma(\tau), \dot\gamma(\tau)) \Big) d\tau \qquad (6)
\end{aligned}$$

and consequently

$$\begin{aligned}
\frac{\partial \Lambda}{\partial t}(t, x, y) &+ \sum_{i=1}^{d} \operatorname{Hess}_{g(\tau_1) \oplus g(\tau_2)} \Lambda \big|_{(t,x,y)} (\xi_i, \xi_i) \\
&\le \frac{d}{\sqrt{t}} \left(\sqrt{\bar\tau_2} - \sqrt{\bar\tau_1}\right) - \frac{1}{2t} \int_{\bar\tau_1 t}^{\bar\tau_2 t} \sqrt{\tau} \left(R_{g(\tau)}(\gamma(\tau)) + |\dot\gamma(\tau)|_{g(\tau)}^2 \right) d\tau \\
&= \frac{d}{\sqrt{t}} \left(\sqrt{\bar\tau_2} - \sqrt{\bar\tau_1}\right) - \frac{1}{2t} \Lambda(t, x, y).
\end{aligned}$$

The proof of Theorem 2 is now achieved under Assumption 1 by combining Lemma 1 and Proposition 1:

Proof of Theorem 2 under Assumption 1. Lemma 6 below ensures that $\Theta(t, \tilde X_t, \tilde Y_t)$ is integrable. Thus it suffices to show that the bounded variation part of $\Theta(t, \tilde X_t, \tilde Y_t)$ is nonpositive. By

Lemma 1,

$$\begin{aligned}
d\Theta(t,\tilde X_t,\tilde Y_t) = \Bigg[&\frac{\partial \Theta}{\partial t}(t,\tilde X_t,\tilde Y_t) \\
&+ \sum_{i=1}^d \operatorname{Hess}_{g(\bar\tau_1 t)\oplus g(\bar\tau_2 t)} \Theta\big|_{(t,\tilde X_t,\tilde Y_t)} \left(\sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e_i^*, \sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e_i^*\right)\Bigg] dt \\
&+ \sum_{i=1}^d \left[\sqrt{2\bar\tau_1}\tilde U_t e_i \oplus \sqrt{2\bar\tau_2}\tilde V_t e_i^*\right]\Theta(t,\tilde X_t,\tilde Y_t) dW_t^i.
\end{aligned}$$

For the bounded variation part we obtain

$$\frac{\partial \Theta}{\partial t}(t,\tilde X_t,\tilde Y_t) = \frac{\sqrt{\bar\tau_2}-\sqrt{\bar\tau_1}}{\sqrt{t}}\Lambda(t,\tilde X_t,\tilde Y_t) + 2\left(\sqrt{\bar\tau_2 t}-\sqrt{\bar\tau_1 t}\right)\frac{\partial \Lambda}{\partial t}(t,\tilde X_t,\tilde Y_t) - 2d\left(\sqrt{\bar\tau_2}-\sqrt{\bar\tau_1}\right)^2$$

and

$$\begin{aligned}
&\sum_{i=1}^d \operatorname{Hess}_{g(\bar\tau_1 t)\oplus g(\bar\tau_2 t)} \Theta\big|_{(t,\tilde X_t,\tilde Y_t)} \left(\sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e_i^*, \sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e_i^*\right) \\
&= 2\left(\sqrt{\bar\tau_2 t}-\sqrt{\bar\tau_1 t}\right)\sum_{i=1}^d \operatorname{Hess}_{g(\bar\tau_1 t)\oplus g(\bar\tau_2 t)} \Lambda\big|_{(t,\tilde X_t,\tilde Y_t)} \left(\sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e_i^*, \sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e_i^*\right).
\end{aligned}$$

Thus, by Proposition 1,

$$\begin{aligned}
&\frac{\partial \Theta}{\partial t}(t,\tilde X_t,\tilde Y_t) + \sum_{i=1}^d \operatorname{Hess}_{g(\bar\tau_1 t)\oplus g(\bar\tau_2 t)} \Theta\big|_{(t,\tilde X_t,\tilde Y_t)} \left(\sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e_i^*, \sqrt{\bar\tau_1}\tilde U_t e_i \oplus \sqrt{\bar\tau_2}\tilde V_t e_i^*\right) \\
&\quad \le 2\left(\sqrt{\bar\tau_2 t}-\sqrt{\bar\tau_1 t}\right)\left[\frac{d}{\sqrt{t}}\left(\sqrt{\bar\tau_2}-\sqrt{\bar\tau_1}\right) - \frac{1}{2t}\Lambda(t,\tilde X_t,\tilde Y_t)\right] \\
&\qquad\qquad + \frac{\sqrt{\bar\tau_2}-\sqrt{\bar\tau_1}}{\sqrt{t}}\Lambda(t,\tilde X_t,\tilde Y_t) - 2d\left(\sqrt{\bar\tau_2}-\sqrt{\bar\tau_1}\right)^2 \\
&\quad = 0.
\end{aligned}$$

Hence $\Theta(t,\tilde X_t,\tilde Y_t)$ is indeed a supermartingale. □

Remark 7. Unlike the case in [1], the pathwise contraction of $\Theta(t,\tilde X_t,\tilde Y_t)$ is no longer true in our case. In other words, the martingale part of $\Theta(t,\tilde X_t,\tilde Y_t)$ does not vanish. We will see it in the following. The minimal $\mathcal{L}$-geodesic $\gamma = \gamma_{xy}^{\tau_1\tau_2}$ of $L(x,\tau_1;y,\tau_2)$ satisfies the $\mathcal{L}$-geodesic equation

$$\nabla^{g(\tau)}_{\dot\gamma(\tau)}\dot\gamma(\tau) = \frac{1}{2}\nabla^{g(\tau)}R_{g(\tau)} - 2\operatorname{Ric}^{\#}_{g(\tau)}(\dot\gamma(\tau)) - \frac{1}{2\tau}\dot\gamma(\tau) \tag{7}$$

(see [5, Corollary 7.19]). Thus the first variation formula (see [5, Lemma 7.15]) yields

$$\sqrt{2\bar\tau_1}\tilde U_t e_i \oplus \sqrt{2\bar\tau_2}\tilde V_t e_i^*\Lambda(t,\tilde X_t,\tilde Y_t) = \sqrt{2t}\bar\tau_2\langle \tilde V_t e_i^*, \dot\gamma(\bar\tau_2 t)\rangle_{g(\bar\tau_2 t)} - \sqrt{2t}\bar\tau_1\langle \tilde U_t e_i, \dot\gamma(\bar\tau_1 t)\rangle_{g(\bar\tau_1 t)}. \tag{8}$$

One obstruction to pathwise contraction is on the difference of time-scalings $\bar\tau_1$ and $\bar\tau_2$. In addition, by (7), $\sqrt{\tau}\dot\gamma(\tau)$ is *not* space-time parallel to γ in general (cf. Remark 5).

4 Proof of Proposition 1

In this section, we write $\tau_1 := \bar{\tau}_1 t$ and $\tau_2 := \bar{\tau}_2 t$. We assume $\tau_2 < T$. For simplicity of notations, we abbreviate the dependency on the metric $g(\tau)$ of several geometric quantities such as Ric, R, the inner product $\langle \cdot, \cdot \rangle$, the covariant derivative ∇ etc. when our choice of τ is obvious. For this abbreviation, we will think that $\gamma(\tau)$ is in $(M, g(\tau))$ and $\dot{\gamma}(\tau)$ is in $(T_{\gamma(\tau)}M, g(\tau))$. Note that, when $\bar{\tau}_1 = 0$, $\lim_{\tau \downarrow \bar{\tau}_1} \sqrt{\tau}\dot{\gamma}(\tau)$ exists while $\lim_{\tau \downarrow 0} |\dot{\gamma}(\tau)| = \infty$. In any case, $\sqrt{\tau}|\dot{\gamma}(\tau)|$ is bounded (see (29)).

We first compute the time derivative of Λ. When $\bar{\tau}_1 > 0$, by [23, Formulas (A.4) and (A.5)] we have

$$\begin{aligned}\frac{\partial L}{\partial \tau_1}(x, \tau_1; y, \tau_2) &= -\sqrt{\tau_1}\left(R_{g(\tau_1)}(x) - |\dot{\gamma}(\tau_1)|^2\right),\\ \frac{\partial L}{\partial \tau_2}(x, \tau_1; y, \tau_2) &= \sqrt{\tau_2}\left(R_{g(\tau_2)}(y) - |\dot{\gamma}(\tau_2)|^2\right),\end{aligned}$$

so that

$$\begin{aligned}\frac{\partial \Lambda}{\partial t}(t, x, y) &= \bar{\tau}_1 \frac{\partial L}{\partial \tau_1}(x, \tau_1; y, \tau_2) + \bar{\tau}_2 \frac{\partial L}{\partial \tau_2}(x, \tau_1; y, \tau_2)\\ &= \frac{1}{t}\left(\tau_2^{3/2}\left(R(\gamma(\tau_2)) - |\dot{\gamma}(\tau_2)|^2\right) - \tau_1^{3/2}\left(R(\gamma(\tau_1)) - |\dot{\gamma}(\tau_1)|^2\right)\right). \end{aligned} \tag{9}$$

Thus the integration-by-parts yields,

$$\begin{aligned}\frac{\partial \Lambda}{\partial t}(t, x, y) &= \frac{3}{2t}\int_{\tau_1}^{\tau_2} \sqrt{\tau}\left(R(\gamma(\tau)) - |\dot{\gamma}(\tau)|^2\right) d\tau\\ &\quad + \frac{1}{t}\int_{\tau_1}^{\tau_2} \tau^{3/2}\Bigg(\frac{\partial R}{\partial \tau}(\gamma(\tau)) + \nabla_{\dot{\gamma}(\tau)} R(\gamma(\tau))\\ &\qquad - 2\langle \nabla_{\dot{\gamma}(\tau)}\dot{\gamma}(\tau), \dot{\gamma}(\tau)\rangle - 2\operatorname{Ric}(\dot{\gamma}(\tau), \dot{\gamma}(\tau))\Bigg) d\tau. \end{aligned} \tag{10}$$

Note that we have

$$\frac{\partial R}{\partial \tau} = -\Delta R - 2\,|\operatorname{Ric}|^2 \tag{11}$$

(see e.g. [22, Proposition 2.5.4]). Since γ satisfies the $\mathcal{L}$-geodesic equation (7), by substituting (7) and (11) into (10), we obtain (5). Note that the derivation of (9) and (10) is still valid even when $\bar{\tau}_1 = 0$ because of the remark at the beginning of this section. Thus (5) holds when $\bar{\tau}_1 = 0$, too.

In order to estimate $\sum_{i=1}^d \operatorname{Hess}_{g(\tau_1)\oplus g(\tau_2)} \Lambda\big|_{(t,x,y)}(\xi_i, \xi_i)$ we begin with the second variation formula for the $\mathcal{L}$-functional:

Lemma 2 (Second variation formula; [5, Lemma 7.37])**.** *Let $\Gamma : (-\varepsilon, \varepsilon) \times [\tau_1, \tau_2] \to M$ be a variation of γ, $S(s, \tau) := \partial_s \Gamma(s, \tau)$, and $Z(\tau) := \partial_s \Gamma(0, \tau)$ the variation field of Γ. Then*

$$\begin{aligned}\frac{d^2}{ds^2}\bigg|_{s=0} \mathcal{L}(\Gamma_s) &= 2\sqrt{\tau}\,\langle \dot{\gamma}(\tau), \nabla_{Z(\tau)} S(0, \tau)\rangle\big|_{\tau=\tau_1}^{\tau=\tau_2} - 2\sqrt{\tau}\operatorname{Ric}(Z(\tau), Z(\tau))\big|_{\tau=\tau_1}^{\tau=\tau_2}\\ &\quad + \frac{1}{\sqrt{\tau}}|Z(\tau)|^2\bigg|_{\tau=\tau_1}^{\tau=\tau_2} - \int_{\tau_1}^{\tau_2} \sqrt{\tau} H(\dot{\gamma}(\tau), Z(\tau)) d\tau\\ &\quad + \int_{\tau_1}^{\tau_2} 2\sqrt{\tau}\left|\nabla_{\dot{\gamma}(\tau)} Z(\tau) + \operatorname{Ric}^{\#}(Z(\tau)) - \frac{1}{2\tau}Z(\tau)\right|^2 d\tau, \end{aligned} \tag{12}$$

where

$$
\begin{aligned}
H(\dot\gamma(\tau), Z(\tau)) := -2\frac{\partial \operatorname{Ric}}{\partial \tau}(Z(\tau), Z(\tau)) &- \operatorname{Hess} R(Z(\tau), Z(\tau)) + 2\left|\operatorname{Ric}^{\#}(Z(\tau))\right|^2 \\
&- \frac{1}{\tau}\operatorname{Ric}(Z(\tau), Z(\tau)) - 2\operatorname{Rm}(Z(\tau), \dot\gamma(\tau), \dot\gamma(\tau), Z(\tau)) \\
&- 4(\nabla_{\dot\gamma(\tau)} \operatorname{Ric})(Z(\tau), Z(\tau)) + 4(\nabla_{Z(\tau)} \operatorname{Ric})(\dot\gamma(\tau), Z(\tau)).
\end{aligned} \tag{13}
$$

In [5] this lemma is only proved in the case $\tau_1 = 0$ and $Z(\tau_1) = 0$. However, the proof given there can be easily adapted to the slightly more general case needed here.

Corollary 1 (see [5, Lemma 7.39] for a similar statement)**.** *If the variation field Z is of the form*

$$
Z(\tau) = \sqrt{\frac{\tau}{t}} Z^*(\tau) \tag{14}
$$

with a space-time parallel field Z^* satisfying $|Z^*(\tau)| \equiv 1$, then

$$
\begin{aligned}
\left.\frac{d^2}{ds^2}\right|_{s=0} \mathcal{L}(\Gamma_s) = 2\sqrt{\tau}\, \langle \dot\gamma(\tau), \nabla_{Z(\tau)} S(0,\tau)\rangle_{g(\tau)}\Big|_{\tau=\tau_1}^{\tau=\tau_2} &- 2\sqrt{\tau}\operatorname{Ric}(Z(\tau), Z(\tau))\Big|_{\tau=\tau_1}^{\tau=\tau_2} \\
- \int_{\tau_1}^{\tau_2} \sqrt{\tau} H(\dot\gamma(\tau), Z(\tau)) d\tau &+ \left.\frac{\sqrt{\tau}}{t}\right|_{\tau=\tau_1}^{\tau=\tau_2}.
\end{aligned}
$$

Proof. Since Z^* is space-time parallel, Z satisfies

$$
\nabla_{\dot\gamma(\tau)} Z(\tau) = -\operatorname{Ric}^{\#}(Z(\tau)) + \frac{1}{2\tau} Z(\tau), \tag{15}
$$

so that the last term in (12) vanishes. □

Corollary 2 (Hessian of L; see [5, Corollary 7.40] for a similar statement)**.** *Let Z be a vector field along γ of the form* (14) *and $\xi := Z(\tau_1) \oplus Z(\tau_2) \in T_{(x,y)}(M\times M)$ Then*

$$
\begin{aligned}
\operatorname{Hess}_{g(\tau_1)\oplus g(\tau_2)} L\big|_{(x,\tau_1;y,\tau_2)}(\xi,\xi) \le -\int_{\tau_1}^{\tau_2} \sqrt{\tau} H(\dot\gamma(\tau), Z(\tau)) d\tau &+ \left.\frac{\sqrt{\tau}}{t}\right|_{\tau=\tau_1}^{\tau=\tau_2} \\
&- 2\sqrt{\tau}\operatorname{Ric}_{g(\tau)}(Z(\tau), Z(\tau))\Big|_{\tau=\tau_1}^{\tau=\tau_2}.
\end{aligned} \tag{16}
$$

Proof. Let $\Gamma : (-\varepsilon, \varepsilon) \times [\tau_1, \tau_2] \to M$ be any variation of γ with variation field Z and such that $\nabla_{Z(\tau_1)} S(0,\tau_1)$ and $\nabla_{Z(\tau_2)} S(0,\tau_2)$ vanish. Since

$$
\operatorname{Hess}_{g(\tau_1)\oplus g(\tau_2)} L\big|_{(x,\tau_1;y,\tau_2)}(\xi,\xi) \le \left.\frac{d^2}{ds^2}\right|_{s=0} \mathcal{L}(\Gamma_s),
$$

the claim follows from Corollary 1. □

Let now Z_i^* $(i = 1, \dots, d)$ be space-time parallel fields along γ satisfying $Z_i^*(\tau_1) = ue_i$ (and consequently $Z_i^*(\tau_2) = ve_i^*$), and $Z_i(\tau) := \sqrt{\tau/t} Z_i^*(\tau)$ (so that $\xi_i = Z_i(\tau_1) \oplus Z_i(\tau_2)$). In order to estimate $\sum_{i=1}^d \operatorname{Hess}_{g(\tau_1)\oplus g(\tau_2)} L\big|_{(x,\tau_1;y,\tau_2)}(\xi_i,\xi_i)$ using Corollary 2 we will compute $\sum_{i=1}^d H(\dot\gamma(\tau), Z_i(\tau))$ in the following (see [5, Section 7.5.3] for a similar argument). Set I_1, I_2

and I_3 by

$$\begin{aligned}
I_1 &:= -2\sum_{i=1}^{d} \frac{\partial \operatorname{Ric}}{\partial \tau}(Z_i(\tau), Z_i(\tau)),\\
I_2 &:= \sum_{i=1}^{d} \Big[-\operatorname{Hess} R(Z_i(\tau), Z_i(\tau)) + 2\,|\operatorname{Ric}^{\#}(Z_i(\tau))|^2 \\
&\qquad\qquad - \frac{1}{\tau}\operatorname{Ric}(Z_i(\tau), Z_i(\tau)) - 2\operatorname{Rm}(Z_i(\tau), \dot\gamma(\tau), \dot\gamma(\tau), Z_i(\tau))\Big],\\
I_3 &:= 4\sum_{i=1}^{d} \Big[(\nabla_{Z_i(\tau)} \operatorname{Ric})(Z_i(\tau), \dot\gamma(\tau)) - (\nabla_{\dot\gamma(\tau)} \operatorname{Ric})(Z_i(\tau), Z_i(\tau))\Big].
\end{aligned}$$

Then $\sum_{i=1}^{d} H(\dot\gamma(\tau), Z_i(\tau)) = I_1 + I_2 + I_3$ holds. By a direct computation,

$$I_2 = \frac{\tau}{t}\left(-\Delta R(\gamma(\tau)) + 2\,|\operatorname{Ric}|^2(\gamma(\tau)) - \frac{1}{\tau} R(\gamma(\tau)) + 2\operatorname{Ric}(\dot\gamma(\tau), \dot\gamma(\tau))\right). \tag{17}$$

The contracted Bianchi identity $\operatorname{div}\operatorname{Ric} = \frac{1}{2}\nabla R$ [13, Lemma 7.7] yields

$$I_3 = \frac{4\tau}{t}\big((\operatorname{div}\operatorname{Ric})(\dot\gamma(\tau)) - (\nabla_{\dot\gamma(\tau)} R)(\gamma(\tau))\big) = -\frac{2\tau}{t}(\nabla_{\dot\gamma(\tau)} R)(\gamma(\tau)). \tag{18}$$

For I_1, we have

$$\begin{aligned}
I_1 &= -2\sum_{i=1}^{d}\left[\frac{d}{d\tau}(\operatorname{Ric}(Z_i(\tau), Z_i(\tau))) - (\nabla_{\dot\gamma(\tau)} \operatorname{Ric})(Z_i(\tau), Z_i(\tau)) - 2\operatorname{Ric}(\nabla_{\dot\gamma(\tau)} Z_i(\tau), Z_i(\tau))\right]\\
&= -2\frac{d}{d\tau}\left(\frac{\tau}{t} R(\gamma(\tau))\right) + 2\frac{\tau}{t}\nabla_{\dot\gamma(\tau)} R(\gamma(\tau)) + 4\sum_{i=1}^{d} \operatorname{Ric}(\nabla_{\dot\gamma(\tau)} Z_i(\tau), Z_i(\tau))\\
&= -\frac{2\tau}{t}\left(\frac{1}{\tau} R(\gamma(\tau)) + \frac{\partial R}{\partial \tau}(\gamma(\tau))\right) + 4\sum_{i=1}^{d} \operatorname{Ric}(\nabla_{\dot\gamma(\tau)} Z_i(\tau), Z_i(\tau)).
\end{aligned} \tag{19}$$

Since Z_i satisfies (15),

$$\begin{aligned}
4\sum_{i=1}^{d} \operatorname{Ric}(\nabla_{\dot\gamma(\tau)} Z_i(\tau), Z_i(\tau)) &= 4\sum_{i=1}^{d} \operatorname{Ric}(-\operatorname{Ric}^{\#}(Z_i(\tau)) + \frac{1}{2\tau} Z_i(\tau), Z_i(\tau))\\
&= -\frac{2\tau}{t}\left(2\,|\operatorname{Ric}|^2(\gamma(\tau)) - \frac{1}{\tau} R(\gamma(\tau))\right).
\end{aligned} \tag{20}$$

By substituting (20) into (19),

$$I_1 = -\frac{2\tau}{t}\left(\frac{\partial R}{\partial \tau}(\gamma(\tau)) + 2\,|\operatorname{Ric}|^2(\gamma(\tau))\right). \tag{21}$$

Hence, by combining (21), (18) and (17),

$$\begin{aligned}
\sum_{i=1}^{d} H(\dot\gamma(\tau), Z_i(\tau)) &= \frac{\tau}{t}\Big(-2\frac{\partial R}{\partial \tau}(\gamma(\tau)) - 2\,|\operatorname{Ric}|^2(\gamma(\tau)) - \Delta R(\gamma(\tau))\\
&\qquad - \frac{1}{\tau} R(\gamma(\tau)) + 2\operatorname{Ric}(\dot\gamma(\tau), \dot\gamma(\tau)) - 2(\nabla_{\dot\gamma(\tau)} R)(\gamma(\tau))\Big).
\end{aligned}$$

Inserting this into (16) we obtain

$$\begin{aligned}
&\sum_{i=1}^{d} \operatorname{Hess}_{g(\tau_1)\oplus g(\tau_2)} L\big|_{(x,\tau_1;y,\tau_2)} (\xi_i, \xi_i) \\
&\quad \le \frac{1}{t}\int_{\tau_1}^{\tau_2} \tau^{3/2}\Big(2\frac{\partial R}{\partial \tau}(\gamma(\tau)) + 2\left|\operatorname{Ric}\right|^2(\gamma(\tau)) + \Delta R(\gamma(\tau)) \\
&\qquad\qquad\qquad + \frac{1}{\tau}R(\gamma(\tau)) - 2\operatorname{Ric}(\dot\gamma(\tau),\dot\gamma(\tau)) + 2(\nabla_{\dot\gamma(\tau)}R)(\gamma(\tau))\Big)d\tau \\
&\qquad + \frac{d\sqrt{\tau}}{t}\bigg|_{\tau=\tau_1}^{\tau=\tau_2} - \frac{2\tau^{3/2}}{t}R(\gamma(\tau))\bigg|_{\tau=\tau_1}^{\tau=\tau_2} \\
&\quad = \frac{d\sqrt{\tau}}{t}\bigg|_{\tau=\tau_1}^{\tau=\tau_2} + \frac{1}{t}\int_{\tau_1}^{\tau_2}\tau^{3/2}\Big(2\left|\operatorname{Ric}\right|^2(\gamma(\tau)) + \Delta R(\gamma(\tau)) - \frac{2}{\tau}R(\gamma(\tau)) - 2\operatorname{Ric}(\dot\gamma(\tau),\dot\gamma(\tau))\Big)d\tau
\end{aligned}$$

which completes the proof of Proposition 1.

5 Coupling via approximation by geodesic random walks

To avoid a technical difficulty coming from singularity of L on the $\mathcal{L}$-cut locus, we provide an alternative way to constructing a coupling of Brownian motions by space-time parallel transport. In this section, we first define a coupling of geometric random walks which approximate $g(\tau)$-Brownian motion. Next, in order to provide a local uniform control of error terms coming from our discretization, we study several estimates of geometric quantities in subsection 5.1. Those are obtained as a small modification of existing arguments in [5, 23, 25]. The $\mathcal{L}$-cut locus is also reviewed and studied there. Finally, we will establish an analogue of arguments in section 3 for coupled geodesic random walks to complete the proof of Theorem 2.

Let us take a family of minimal $\mathcal{L}$-geodesics $\{\gamma_{xy}^{\tau_1\tau_2} \mid \bar\tau_1 \le \tau_1 < \tau_2 \le \bar\tau_2, x, y \in M\}$ so that a map $(x,\tau_1;y,\tau_2) \mapsto \gamma_{xy}^{\tau_1\tau_2}$ is measurable. The existence of such a family of minimal $\mathcal{L}$-geodesics can be shown in a similar way as discussed in the proof of [15, Proposition 2.6] since the family of minimal $\mathcal{L}$-geodesics with fixed endpoints is compact (cf. [5, the proof of Lemma 7.27]). For each $\tau \in [\bar\tau_1, T]$, take a measurable section $\Phi^{(\tau)}$ of $g(\tau)$-orthonormal frame bundle $\mathcal{O}^{g(\tau)}(M)$ of M. For $x, y \in M$ and $\tau_1, \tau_2 \in [\bar\tau_1, T]$ with $\tau_1 < \tau_2$, let us define $\Phi_i(x,\tau_1;y,\tau_2) \in \mathcal{F}(M)$ for $i = 1, 2$ by

$$\begin{aligned}
\Phi_1(x,\tau_1;y,\tau_2) &:= \Phi^{(\tau_1)}(x), \\
\Phi_2(x,\tau_1;y,\tau_2) &:= m_{xy}^{\tau_1\tau_2} \circ \Phi^{(\tau_1)}(x),
\end{aligned}$$

where $m_{xy}^{\tau_1\tau_2}$ is as given in Definition 2. Let us take a family of $\mathbb{R}^d$-valued i.i.d. random variables $(\lambda_n)_{n\in\mathbb{N}}$ which are uniformly distributed on a unit ball centered at origin. We denote the (Riemannian) exponential map with respect to $g(\tau)$ at $x \in M$ by $\exp_x^{(\tau)}$. In what follows, we define a coupled geodesic random walk $\mathbf{X}_t^\varepsilon = (X_{\bar\tau_1 t}^\varepsilon, Y_{\bar\tau_2 t}^\varepsilon)$ with scale parameter $\varepsilon > 0$ and initial condition $\mathbf{X}_s^\varepsilon = (x_1, y_1)$ inductively. First we set $(X_{\bar\tau_1 s}^\varepsilon, Y_{\bar\tau_2 s}^\varepsilon) := (x_1, y_1)$. For simplicity of notations, we set $t_n := (s + \varepsilon^2 n) \wedge (T/\bar\tau_2)$. After we defined $(\mathbf{X}_t^\varepsilon)_{t\in[s,t_n]}$, we extend it to $(\mathbf{X}_t^\varepsilon)_{t\in[s,t_{n+1}]}$ by

$$\begin{aligned}
\hat\lambda_{n+1}^{(i)} &:= \sqrt{d+2}\,\Phi_i(X_{\bar\tau_1 t_n}^\varepsilon, \bar\tau_1 t_n; Y_{\bar\tau_2 t_n}^\varepsilon, \bar\tau_2 t_n)\lambda_{n+1}, \qquad i = 1, 2, \\
X_{\bar\tau_1 t}^\varepsilon &:= \exp_{X_{\bar\tau_1 t_n}^\varepsilon}^{(\bar\tau_1 t_n)}\left(\frac{t - t_n}{\varepsilon}\sqrt{2\bar\tau_1}\hat\lambda_{n+1}^{(1)}\right), \\
Y_{\bar\tau_2 t}^\varepsilon &:= \exp_{Y_{\bar\tau_2 t_n}^\varepsilon}^{(\bar\tau_2 t_n)}\left(\frac{t - t_n}{\varepsilon}\sqrt{2\bar\tau_2}\hat\lambda_{n+1}^{(2)}\right).
\end{aligned}$$

We can (and we will) extend the definition of X^ε_τ for $\tau \in [T\bar\tau_1/\bar\tau_2, T]$ in the same way. As in section 3, $X^\varepsilon_{\bar\tau_1 t}$ does not move when $\bar\tau_1 = 0$. Note that $\sqrt{d+2}$ is a normalization factor in the sense $\mathrm{Cov}(\sqrt{d+2}\lambda_n) = \mathrm{Id}$. Let us equip path spaces $C([a,b] \to M)$ or $C([a,b] \to M \times M)$ with the uniform convergence topology induced from $g(T)$. Here the interval $[a,b]$ will be chosen appropriately in each context. By (23) which we will review below, different choices of a metric $g(\tau)$ from $g(T)$ always induce the same topology on path spaces. As shown in [11], $(X^\varepsilon_\tau)_{\tau\in[\bar\tau_1 s,T]}$ and $(Y^\varepsilon_\tau)_{\tau\in[\bar\tau_2 s,T]}$ converge in law to $g(\tau)$-Brownian motions $(X_\tau)_{\tau\in[\bar\tau_1 s,T]}$ and $(Y_\tau)_{\tau\in[\bar\tau_2 s,T]}$ on M with initial conditions $X_{\bar\tau_1 s} = x_1$, $Y_{\bar\tau_2 s} = y_1$ respectively as $\varepsilon \to 0$ (when $\bar\tau_1 > 0$). As a result, $\mathbf{X}^\varepsilon$ is tight and hence there is a convergent subsequence of $\mathbf{X}^\varepsilon$. We fix such a subsequence and use the same symbol $(\mathbf{X}^\varepsilon)_\varepsilon$ for simplicity of notations. We denote the limit in law of $\mathbf{X}^\varepsilon$ as $\varepsilon \to 0$ by $\mathbf{X}_t = (X_{\bar\tau_1 t}, Y_{\bar\tau_2 t})$. Recall that, in this paper, $g(\tau)$-Brownian motion means a time-inhomogeneous diffusion process associated with $\Delta_{g(\tau)}$ instead of $\Delta_{g(\tau)}/2$.

Remark 8. We explain the reason why our alternative construction works efficiently to avoid the problem arising from singularity of L. To make it clear, we begin with observing the essence of difficulties in the SDE approach we used in section 3. Recall that our argument is based on the Itô formula. Hence non-differentiability of L at the $\mathcal{L}$-cut locus causes the technical difficulty. One possible strategy is to extend the Itô formula for $\mathcal{L}$-distance. Since $\mathcal{L}$-cut locus is sufficiently thin, we can expect that the totality of times when our coupled particles stay there has measure zero. In addition, as that of Riemannian cut locus, the presence of $\mathcal{L}$-cut locus would work to decrease the $\mathcal{L}$-distance between coupled particles. Thus one might think it possible to extend Itô formula for $\mathcal{L}$-distance to the one involving a "local time at the $\mathcal{L}$-cut locus". If we succeed in doing so, we will obtain a differential inequality which implies the supermartingale property by neglecting this additional term since it should be nonpositive.

Instead of completing the above strategy, our alternative approach in this section directly provides a difference inequality without extracting the additional "local time" term. By dividing a minimal $\mathcal{L}$-geodesic into two pieces, we can obtain a "difference inequality" of $\mathcal{L}$-distance even when the pair of endpoints belongs to the $\mathcal{L}$-cut locus (see Lemma 3). In order to employ such an inequality, it is more suitable to work with discrete time processes.

5.1 Preliminaries on the geometry of $\mathcal{L}$-functional

Recall that we assume that M has bounded curvature, so that there is a constant $C_0 < \infty$ such that

$$\max_{(x,\tau)\in M\times[\bar\tau_1,T]} |\operatorname{Rm}|_{g(\tau)}(x) \vee |\operatorname{Ric}|_{g(\tau)}(x) \le C_0. \tag{22}$$

On the basis of (22), we have a comparison of Riemannian metrics at different times. That is, for $\tau_1 < \tau_2$,

$$\mathrm{e}^{-2C_0(\tau_2-\tau_1)} g(\tau_2) \le g(\tau_1) \le \mathrm{e}^{2C_0(\tau_2-\tau_1)} g(\tau_2). \tag{23}$$

Let $\rho_{g(\tau)}$ be the distance function on M at time τ. Note that a similar comparison between $\rho_{g(\tau_1)}$ and $\rho_{g(\tau_2)}$ follows from (23). We also obtain the following bounds for L from (22) and (23). Let $\gamma : [\tau_1, \tau_2] \to M$ be a minimal $\mathcal{L}$-geodesic. Then, for $\tau \in [\tau_1, \tau_2]$,

$$\begin{aligned} \frac{\mathrm{e}^{-2C_0(\tau_2-\tau_1)}}{2\sqrt{\tau_2}-\sqrt{\tau_1}} \rho_{g(T)}(\gamma(\tau_1), \gamma(\tau))^2 - \frac{2}{3} dC_0(\tau_2^{3/2} - \tau_1^{3/2}) &\le L(\gamma(\tau_1), \tau_1; \gamma(\tau_2), \tau_2) \\ &\le \frac{\mathrm{e}^{2C_0(\tau_2-\tau_1)}}{2\sqrt{\tau_2}-\sqrt{\tau_1}} \rho_{g(T)}(\gamma(\tau_1), \gamma(\tau_2))^2 + \frac{2}{3} dC_0(\tau_2^{3/2} - \tau_1^{3/2}) \end{aligned} \tag{24}$$

(see [5, Lemma 7.13] and [23, Proposition B.2]). The same estimate holds for $\rho_{g(T)}(\gamma(\tau), \gamma(\tau_2))^2$ instead of $\rho_{g(T)}(\gamma(\tau_1), \gamma(\tau))^2$. Taking the fact that $\mathcal{L}$-functional is *not* invariant under reparametrization of curves into account, we will introduce an estimate for the velocity of the

minimal $\mathcal{L}$-geodesic γ. By a similar argument as in [5, Lemma 7.13 (ii)], there exists $\tau^* \in [\tau_1, \tau_2]$ such that

$$\tau^* \left|\dot\gamma(\tau^*)\right|^2_{g(\tau^*)} \le \frac{1}{2(\sqrt{\tau_2} - \sqrt{\tau_1})} \left(L(\gamma(\tau_1), \tau_1; \gamma(\tau_2), \tau_2) + \frac{2dC_0}{3}(\tau_2^{3/2} - \tau_1^{3/2}) \right). \tag{25}$$

Suppose $\tau_2 < T$. Then, as shown in [21] (see [6] also), (22) yields that there is a constant $C(d) > 0$ depending only on d such that

$$\sup_{\tau \le \tau_2,\, x \in M} \left|\nabla \mathrm{Rm}\right|_{g(\tau)}(x) \le \frac{C(d)C_0}{(T - \tau_2) \wedge C_0^{-1}} =: C_0'. \tag{26}$$

By virtue of (26), there exists a constant $c_1, C_1 > 0$ which depends on C_0, C_0' and T such that for all $\tau_1', \tau_2' \in [\tau_1, \tau_2]$ with $\tau_1' < \tau_2'$,

$$\tau_2' \left|\dot\gamma(\tau_2')\right|^2_{g(\tau_2')} \le c_1 \tau_1' \left|\dot\gamma(\tau_1')\right|^2_{g(\tau_1')} + C_1, \tag{27}$$

$$\tau_1' \left|\dot\gamma(\tau_1')\right|^2_{g(\tau_1')} \le c_1 \tau_2' \left|\dot\gamma(\tau_2')\right|^2_{g(\tau_2')} + C_1. \tag{28}$$

The first inequality in (27) can be shown similarly as [5, Lemma 7.24]. It is due to a differential inequality based on the $\mathcal{L}$-geodesic equation (7) which provides an upper bound of $\partial_\tau(\tau|\dot\gamma(\tau)|^2_{g(\tau)})$. By considering a lower bound of the same quantity instead, we obtain the second inequality (28) in a similar way. Combining (27) and (28) with (25) and (24), we can take constants $c_2 > 0$ and $C_2 > 0$ depending on C_0, c_1, C_1, τ_1 and τ_2 such that

$$\tau \left|\dot\gamma(\tau)\right|^2_{g(\tau)} \le c_2 \rho_{g(T)}(\gamma(\tau_1), \gamma(\tau_2))^2 + C_2 \tag{29}$$

for $\tau_1 \le \tau \le \tau_2$. Though c_2 and C_2 depends on τ_1 and τ_2, it is easy to see that c_2 and C_2 are uniformly bounded above as long as $\tau_2 - \tau_1$ and $T - \tau_2$ is uniformly away from 0.

Let us recall the definition and some properties of $\mathcal{L}$-cut locus according to [5, 23, 25]. Given $\tau, \tau' \in [0, T)$ with $\tau < \tau'$ and $x \in M$, we define $\mathcal{L}$-exponential map $\mathcal{L}_{\tau,\tau'} \exp_x : T_xM \to M$ by $\mathcal{L}_{\tau,\tau'} \exp_x(Z) = \gamma(\tau')$, where γ is a unique $\mathcal{L}$-geodesic from (τ, x) with the initial condition $\lim_{\tau' \downarrow \tau} \sqrt{\tau'}\dot\gamma(\tau') = Z$. Note that we can extend the domain of $\mathcal{L}$-geodesic γ to the interval $[\tau, T)$ by using (27) (see [5, Lemma 7.25]). Set

$$\Omega(x, \tau_1; \tau_2) := \left\{ Z \in T_xM \;\middle|\; \begin{array}{l} \gamma \,:\, [\tau_1, \tau_2] \to M \text{ given by } \gamma(\tau) := \mathcal{L}_{\tau_1,\tau} \exp_x(Z) \\ \text{is a minimal } \mathcal{L}\text{-geodesic} \end{array} \right\},$$

$$\bar\tau(x, \tau; Z) := \sup\left\{ \tau \in (\tau_1, T) \mid Z \in \Omega(x, \tau_1; \tau) \right\}.$$

Based on these notations, we define the $\mathcal{L}$-cut locus $\mathcal{L}\mathrm{Cut}$ by

$$\mathcal{L}\mathrm{Cut} := \left\{ (x, \tau_1; y, \tau_2) \;\middle|\; \begin{array}{l} x \in M, \tau_1 \in [\bar\tau_1, T), \\ y = \mathcal{L}_{\tau_1,\tau_2} \exp_x(Z) \text{ for some } Z \in T_xM, \\ \tau_2 = \bar\tau(x, \tau_1; Z) \in (\tau_1, T) \end{array} \right\}.$$

As remarked in [5, 23, 25], $\mathcal{L}\mathrm{Cut}$ is a union of two different kinds of sets. The first one consists of $(x, \tau_1; y, \tau_2)$ such that there exists more than one minimal $\mathcal{L}$-geodesics joining (x, τ_1) and (y, τ_2). The second consists of $(x, \tau_1; y, \tau_2)$ such that (y, τ_2) is conjugate to (x, τ_1) along a minimal $\mathcal{L}$-geodesic with respect to $\mathcal{L}$-Jacobi field. Note that we can define exponential map in the reverse direction in τ. By using this reverse exponential map, "reversed $\mathcal{L}$-cut locus" is defined and it is identified with $\mathcal{L}\mathrm{Cut}$ by virtue of the above characterization of $\mathcal{L}\mathrm{Cut}$.

5.2 Proof of Theorem 2

For the proof of our main theorem based on discrete approximation, we will follow a similar way as in previous studies in this direction (see [10, 11] and references therein). Our first task is to show a difference inequality of $\Lambda(t, \mathbf{X}_t^\varepsilon)$ in Lemma 3. We begin with introducing some notations. Set $\gamma_n := \gamma^{\bar{\tau}_1 t_n, \bar{\tau}_2 t_n}_{\mathbf{X}^\varepsilon_{t_n}}$ and let us define a vector field $\hat{\lambda}^\dagger_{n+1}$ along γ_n by $\hat{\lambda}_{n+1}(\tau) = \sqrt{\tau/t_n}\lambda^*_{n+1}(\tau)$, where λ^*_{n+1} is a space-time parallel vector field along γ_n with initial condition $\hat{\lambda}^*_{n+1}(\bar{\tau}_1 t_n) = \hat{\lambda}^{(1)}_{n+1}$. Let us define random variables ζ_n and Σ_n as follows:

$$
\begin{aligned}
\zeta_{n+1} &:= \sqrt{2\tau}\ \langle \hat{\lambda}^\dagger_{n+1}(\tau), \dot{\gamma}_n(\tau)\rangle_{g(\tau)}\Big|_{\tau=\bar{\tau}_1 t_n}^{\bar{\tau}_2 t_n},\\
\Sigma_{n+1} &:= \frac{1}{t_n}\,\tau^{3/2}\left(R_{g(\tau)}(\gamma_n(\tau)) - |\dot{\gamma}_n(\tau)|^2_{g(\tau)}\right)\Big|_{\tau=\bar{\tau}_1 t_n}^{\bar{\tau}_2 t_n}\\
&\qquad + \left(\left(\frac{\sqrt{\tau}}{t_n} - 2\sqrt{\tau}\,\mathrm{Ric}_{g(\tau)}(\hat{\lambda}^\dagger_{n+1}(\tau), \hat{\lambda}^\dagger_{n+1}(\tau))\right)\Big|_{\tau=\bar{\tau}_1 t_n}^{\bar{\tau}_2 t_n} \right.\\
&\qquad\qquad \left. - \int_{\bar{\tau}_1 t_n}^{\bar{\tau}_2 t_n} \sqrt{\tau} H\left(\dot{\gamma}(\tau), \hat{\lambda}^\dagger_{n+1}(\tau)\right) d\tau \right).
\end{aligned}
$$

Here H is as given in (13). For $M_0 \subset M$, we define $\sigma_{M_0} : C([s, T/\bar{\tau}_2] \to M \times M) \to [0,\infty)$ by

$$\sigma_{M_0}(w, \tilde{w}) := \inf\{t \geq s \mid w_{\bar{\tau}_1 t} \notin M_0 \text{ or } w_{\bar{\tau}_2 t} \notin M_0\}.$$

For simplicity of notations, $\sigma_{M_0}(\mathbf{X}^\varepsilon)$ and $\sigma_{M_0}(\mathbf{X})$ are denoted by $\sigma^\varepsilon_{M_0}$ and $\sigma^0_{M_0}$ respectively. As shown in [11], for any $\eta > 0$, we can take a compact set $M_0 \subset M$ such that $\lim_{\varepsilon\to 0}\mathbb{P}[\sigma^\varepsilon_{M_0} \leq T] \leq \eta$ holds (cf. [12]).

Lemma 3. *Let $M_0 \subset M$ be compact. Then there exist a family of random variables $(Q^\varepsilon_n)_{n\in\mathbb{N},\varepsilon>0}$ and a family of deterministic constants $(\delta(\varepsilon))_{\varepsilon>0}$ with $\lim_{\varepsilon\to 0}\delta(\varepsilon) = 0$ satisfying*

$$\sum_{n;\, t_n < \sigma^\varepsilon_{M_0}\wedge(T/\bar{\tau}_2)} Q^\varepsilon_n \leq \delta(\varepsilon) \tag{30}$$

such that

$$\Lambda(t_{n+1}, \mathbf{X}^\varepsilon_{t_{n+1}}) \leq \Lambda(t_n, \mathbf{X}^\varepsilon_{t_n}) + \varepsilon\zeta_{n+1} + \varepsilon^2\Sigma_{n+1} + Q^\varepsilon_{n+1}. \tag{31}$$

Proof. When $(X^\varepsilon(\bar{\tau}_1 t_n), \bar{\tau}_1 t_n; Y^\varepsilon(\bar{\tau}_2 t_n), \bar{\tau}_2 t_n) \notin \mathcal{L}\mathrm{Cut}$, the inequality (31) follows from the Taylor expansion with the error term $Q^\varepsilon_{n+1} = o(\varepsilon^2)$. Indeed, the first variation formula ([5, Lemma 7.15] cf. (8)) produces $\varepsilon\zeta_{n+1}$ and Corollary 2 together with (9) implies the bound $\varepsilon^2\Sigma_{n+1}$ of the second order term. To include the case $(X^\varepsilon(\bar{\tau}_1 t_n), \bar{\tau}_1 t_n; Y^\varepsilon(\bar{\tau}_2 t_n), \bar{\tau}_2 t_n) \in \mathcal{L}\mathrm{Cut}$ as well as to obtain a uniform bound (30), we extend this argument. Set $\tau^*_n := (\bar{\tau}_1 + \bar{\tau}_2)t_n/2$. Then we can show

$$
\begin{aligned}
&(X^\varepsilon_{\bar{\tau}_1 t_n}, \bar{\tau}_1 t_n; \gamma_n(\tau^*_n), \tau^*_n) \notin \mathcal{L}\mathrm{Cut},\\
&(\gamma_n(\tau^*_n), \bar{\tau}^*_n; X^\varepsilon_{\bar{\tau}_2 t_n}, \bar{\tau}_2 t_n) \notin \mathcal{L}\mathrm{Cut}
\end{aligned}
$$

since minimal $\mathcal{L}$-geodesics with these pair of endpoints can be extended with keeping its minimality (cf. see [5, Section 7.8] and [25]). Set $x^*_{n+1} = \exp^{(\tau^*_n)}_{\gamma_n(\tau^*_n)}\left(\sqrt{\bar{\tau}_1 + \bar{\tau}_2}\lambda^\dagger_{n+1}(\tau^*_n)\right)$. The triangle inequality for L yields

$$
\begin{aligned}
\Lambda(t_n, \mathbf{X}^\varepsilon_{t_n}) &= L(X^\varepsilon_{\bar{\tau}_1 t_n}, \bar{\tau}_1 t_n; \gamma_n(\tau^*_n), \tau^*_n) + L(\gamma_n(\tau^*_n), \tau^*_n; X^\varepsilon_{\bar{\tau}_2 t_n}, \bar{\tau}_2 t_n),\\
\Lambda(t_{n+1}, \mathbf{X}^\varepsilon_{t_{n+1}}) &\leq L\left(X^\varepsilon_{\bar{\tau}_1 t_{n+1}}, \bar{\tau}_1 t_{n+1}; x^*_{n+1}, \tau^*_{n+1}\right) + L\left(x^*_{n+1}, \tau^*_{n+1}; X^\varepsilon_{\bar{\tau}_2 t_{n+1}}, \bar{\tau}_2 t_{n+1}\right).
\end{aligned}
$$

Hence

$$\Lambda(t_{n+1}, \mathbf{X}^\varepsilon_{t_{n+1}}) - \Lambda(t_n, \mathbf{X}^\varepsilon_{t_n}) \le \left(L\left(X^\varepsilon_{\bar\tau_1 t_{n+1}}, \bar\tau_1 t_{n+1}; x^*_{n+1}, \tau^*_{n+1} \right) - L(X^\varepsilon_{\bar\tau_1 t_n}, \bar\tau_1 t_n; \gamma_n(\tau^*_n), \tau^*_n) \right) + \left(L\left(x^*_{n+1}, \tau^*_{n+1}; X^\varepsilon_{\bar\tau_2 t_{n+1}}, \bar\tau_2 t_{n+1} \right) - L(\gamma_n(\tau^*_n), \tau^*_n; X^\varepsilon_{\bar\tau_2 t_n}, \bar\tau_2 t_n) \right)$$

and the desired inequality with $Q^\varepsilon_n = o(\varepsilon^2)$ holds by applying the Taylor expansion to each term on the right hand side of the above inequality.

We turn to showing the claimed control (30) of the error term Q^ε_n. Take $M_1 \supset M_0$ compact such that every minimal $\mathcal{L}$-geodesic joining $(x, \bar\tau_1 t)$ and $(y, \bar\tau_2 t)$ is included in M_1 if $x, y \in M_0$ and $t \in [s, T/\bar\tau_2]$. Indeed, such M_1 exists since we have the lower bound of L in (24) and L is continuous. Let us define a set A by

$$A := \left\{ ((\tau_1, x), (\tau_3, z), (\tau_2, y)) \in ([\bar\tau_1, T] \times M_1)^3 \;\middle|\; \begin{array}{l} x, y \in M_0, \\ \tau_2 - \tau_1 \ge (\bar\tau_2 - \bar\tau_1)s, \\ \tau_3 = (\tau_1 + \tau_2)/2, \\ L(x, \tau_1; z, \tau_3) + L(z, \tau_3; y, \tau_2) \\ = L(x, \tau_1; y, \tau_2) \end{array} \right\}.$$

Note that A is compact. Let $\pi_1, \pi_2 : A \to ([\bar\tau_1, T] \times M_1)^2$ be defined by

$$\pi_1((\tau_1, x), (\tau_3, z), (\tau_2, y)) := ((\tau_1, x), (\tau_3, z)),$$
$$\pi_2((\tau_1, x), (\tau_3, z), (\tau_2, y)) := ((\tau_3, z), (\tau_2, y)).$$

Then $\pi_1(A)$ and $\pi_2(A)$ are compact and $\pi_i(A) \cap \mathcal{L}\mathrm{Cut} = \emptyset$ for $i = 1, 2$. The second assertion comes from the fact that (z, τ_3) is on a minimal $\mathcal{L}$-geodesic joining (x, τ_1) and (y, τ_2) for $((x, \tau_1), (z, \tau_3), (y, \tau_2)) \in A$. Note that $\mathcal{L}\mathrm{Cut}$ is closed (see [23]; though they assumed M to be compact, an extension to the non-compact case is straightforward). Thus we can take relatively compact open sets $G_1, G_2 \subset [\bar\tau_1, T] \times M$ such that $\pi_i(H) \subset G_i$ and $\bar G_i \cap \mathcal{L}\mathrm{Cut} = \emptyset$ for $i = 1, 2$. Then the Taylor expansion we discussed above can be done on G_1 or G_2 for sufficiently small ε. Recall that L is smooth outside of $\mathcal{L}\mathrm{Cut}$ (see [5]). Thus the convergence $\varepsilon^{-2} Q_n(\varepsilon) \to 0$ as $\varepsilon \to 0$ is uniform in n and independent of $\mathbf{X}^\varepsilon_{t_n}$ as long as $t_n < \sigma^\varepsilon_{M_0} \wedge (T/\bar\tau_2)$. Since the cardinality of $\{ n \mid t_n < \sigma^\varepsilon_{M_0} \wedge (T/\bar\tau_2) \}$ is of order at most ε^{-2}, the assertion (30) holds. □

We next establish the corresponding difference inequality for $\Theta(t, \mathbf{X}^\varepsilon_t)$ (Corollary 3). For that, we show the following auxiliary lemma.

Lemma 4. *Let $T_0 < T$.*

(i) There exist deterministic constants $c'_2 > 0$ and $C'_2 > 0$ such that

$$|\zeta_n| \le c'_2 \rho_{g(T)}(\mathbf{X}^\varepsilon_{t_{n-1}}) + C'_2, \qquad |\Lambda(t_n, \mathbf{X}^\varepsilon_{t_n})| \le c'_2 \rho_{g(T)}(\mathbf{X}^\varepsilon_{t_n})^2 + C'_2$$

if $t_n \le T_0/\bar\tau_2$.

(ii) Take $M_0 \subset M$ compact. Then there is a constant $R = R(T_0, M_0) > 0$ such that $|\Sigma_n| \le R$ holds if $t_n < \sigma^\varepsilon_{M_0} \wedge (T_0/\bar\tau_2)$.

Proof. By the definition of ζ_n, we have

$$|\zeta_n| \le \sqrt{2(d+2)}\, t_{n-1} \left(\bar\tau_1 |\dot\gamma_{n-1}(\bar\tau_1 t_{n-1})|_{g(\bar\tau_1 t_{n-1})} + \bar\tau_2 |\dot\gamma_{n-1}(\tau_2 t_{n-1})|_{g(\bar\tau_2 t_{n-1})} \right).$$

Thus the desired bound for $|\zeta_n|$ follows from (29) and (23). Similarly, the estimate for $\Lambda(t_n, \mathbf{X}^\varepsilon_{t_n})$ follows from (24) and (23). For the assertion (ii), we deal with the integral involving H in the

definition of Σ_n. Note that every tensor field appeared in the definition of H is continuous. As in the proof of Lemma 3, take $M_1 \supset M_0$ compact such that every minimal $\mathcal{L}$-geodesic joining $(x, \bar{\tau}_1 t)$ and $(y, \bar{\tau}_2 t)$ is included in M_1 if $x, y \in M_0$ and $t \in [s, T/\bar{\tau}_2]$. Since $\mathbf{X}^\varepsilon_{t_{n-1}} \in M_0 \times M_0$ holds on the event $\{t_n < \sigma^\varepsilon_{M_0} \wedge (T_0/\bar{\tau}_2)\}$, the upper bound (29) of $\sqrt{\tau}|\dot{\gamma}(\tau)|$ implies that $H(\dot{\gamma}_n(\tau), Z(\tau))$ is uniformly bounded for any vector field $Z(\tau)$ along γ_n of the form $Z(\tau) = \sqrt{\tau/t_n} Z^*(\tau)$ with a space-time parallel vector field $Z^*(\tau)$ satisfying $|Z^*(\tau)|_{g(\tau)} \le 1$.

This fact yields an expected bound for the integral. For any other terms in the definition of Σ_n, we can estimate them as in the assertion (i). □

By virtue of Lemma 4, $\Lambda(t_n, \mathbf{X}^\varepsilon_{t_n})$, ζ_n and Σ_n are uniformly bounded on the event $\{t_n < \sigma^\varepsilon_{M_0} \wedge (T_0/\bar{\tau}_2)\}$ for $T_0 < T$. Thus Lemma 3 yields the following:

Corollary 3. *Let $T_0 < T$ and $M_0 \subset M$ be a compact set. Then there exist a family of random variables $(\tilde{Q}^\varepsilon_n)_{n \in \mathbb{N}, \varepsilon > 0}$ and a family of deterministic constants $(\tilde{\delta}(\varepsilon))_{\varepsilon > 0}$ with $\lim_{\varepsilon \to 0} \tilde{\delta}(\varepsilon) = 0$ satisfying*

$$\sum_{n;\, t_n < \sigma^\varepsilon_{M_0} \wedge (T_0/\bar{\tau}_2)} \tilde{Q}^\varepsilon_n \le \tilde{\delta}(\varepsilon)$$

such that

$$\begin{aligned} \Theta(t_{n+1}, \mathbf{X}^\varepsilon_{t_{n+1}}) \le \Theta(t_n, \mathbf{X}^\varepsilon_{t_n}) &+ \frac{\varepsilon^2}{\sqrt{t_n}} \left(\sqrt{\bar{\tau}_2} - \sqrt{\bar{\tau}_1} \right) \Lambda(t_n, \mathbf{X}^\varepsilon_{t_n}) - 2\varepsilon^2 d \left(\sqrt{\bar{\tau}_2} - \sqrt{\bar{\tau}_1} \right)^2 \\ &+ 2\varepsilon \left(\sqrt{\bar{\tau}_2 t_{n+1}} - \sqrt{\bar{\tau}_1 t_{n+1}} \right) \zeta_{n+1} + 2\varepsilon^2 \left(\sqrt{\bar{\tau}_2 t_{n+1}} - \sqrt{\bar{\tau}_1 t_{n+1}} \right) \Sigma_{n+1} \\ &+ \tilde{Q}^\varepsilon_{n+1}. \end{aligned} \tag{32}$$

The term Σ_n corresponds to the one dominating the bounded variation part of $d\Lambda(t, \tilde{X}_t, \tilde{Y}_t)$ in section 3. However, as a result of our discretization, we are no longer able to apply Proposition 1 directly to estimate Σ_n itself. In this case, we can do it to the conditional expectation of Σ_n instead. Set $\mathcal{G}_n := \sigma(\lambda_1, \dots, \lambda_n)$ and $\bar{\Sigma}_{n+1} := \mathbb{E}[\Sigma_{n+1} \,|\, \mathcal{G}_n]$. Then, since each Φ_i is isometry and $(d+2)\mathbb{E}[\langle \lambda_n, e_i \rangle \langle \lambda_n, e_j \rangle] = \delta_{ij}$, Proposition 1 yields

$$\bar{\Sigma}_n \le \frac{d}{\sqrt{t_n}} \left(\sqrt{\bar{\tau}_2} - \sqrt{\bar{\tau}_1} \right) - \frac{1}{2t_n} \Lambda(t_n, \mathbf{X}^\varepsilon_{t_n}). \tag{33}$$

In order to replace Σ_n with $\bar{\Sigma}_n$ in (32), we show the following:

Lemma 5. *Let $t_0 < T_0 < T$ and $M_0 \subset M$ compact. For $t \in [\bar{\tau}_1, T]$, set $N^\varepsilon_t := \sup\{n \in \mathbb{N} \mid \bar{\tau}_2(s + \varepsilon^2 n) \le t\}$. Then, for $\eta > 0$,*

$$\limsup_{\varepsilon \to 0} \mathbb{P} \left[\sup_{\substack{N^\varepsilon_{t_0} \le N \le N^\varepsilon_{T_0} \\ t_N \le \sigma^\varepsilon_{M_0}}} \left| \sum_{n=1}^N \sqrt{t_n} (\Sigma_n - \bar{\Sigma}_n) \right| > \varepsilon^{-2} \eta \right] = 0.$$

Proof. Lemma 4 ensures that Σ_n and $\bar{\Sigma}_n$ is bounded as long as $n \le N^\varepsilon_{T_0}$ and $t_n < \sigma^\varepsilon_{M_0}$. Thus $\sum_{n=1}^N \sqrt{t_n}(\Sigma_n - \bar{\Sigma}_n)$ is a $\mathcal{G}_N$-local martingale. Hence the Doob inequality implies

$$\limsup_{\varepsilon \to 0} \mathbb{P} \left[\sup_{\substack{0 \le N \le N^\varepsilon_t \\ t_N \le \sigma^\varepsilon_{M_0}}} \left| \sum_{n=1}^N \sqrt{t_n} (\Sigma_n - \bar{\Sigma}_n) \right| > \varepsilon^{-2} \eta \right] = 0 \tag{34}$$

for $t \in [\bar{\tau}_1, T_0]$. Since we have

$$\left\{ \sup_{\substack{N^\varepsilon_{t_0} \le N \le N^\varepsilon_{T_0} \\ t_N \le \sigma^\varepsilon_{M_0}}} \left| \sum_{n=N^\varepsilon_{t_0}}^{N} \sqrt{t_n}(\Sigma_n - \bar{\Sigma}_n) \right| > \varepsilon^{-2}\eta \right\}$$
$$\subset \left\{ \sup_{\substack{0 \le N \le N^\varepsilon_{t_0} \\ t_N \le \sigma^\varepsilon_{M_0}}} \left| \sum_{n=1}^{N} \sqrt{t_n}(\Sigma_n - \bar{\Sigma}_n) \right| > \frac{\varepsilon^{-2}\eta}{2} \right\} \cup \left\{ \sup_{\substack{0 \le N \le N^\varepsilon_{T_0} \\ t_N \le \sigma^\varepsilon_{M_0}}} \left| \sum_{n=1}^{N} \sqrt{t_n}(\Sigma_n - \bar{\Sigma}_n) \right| > \frac{\varepsilon^{-2}\eta}{2} \right\},$$

the assertion follows from (34). □

As a final preparation, we show the following auxiliary lemma.

Lemma 6. *There exists $C_3 > 0$ such that*

$$\mathbb{E}\left[\sup_{s \le t \le T/\bar{\tau}_2} |\Theta(t, \mathbf{X}_t)|^2 \right] < C_3.$$

Proof. By virtue of (24), $\Theta(t, \mathbf{X}_t)$ is bounded from below uniformly in $t \in [s, T/\bar{\tau}_2]$. In addition, there is a constant $c, C > 0$ such that

$$\Theta(t, \mathbf{X}_t) \le c\rho_{g(T)}(\mathbf{X}_t)^2 + C$$

holds for $t \in [s, T/\bar{\tau}_2]$. Take a reference point $o \in M$. Then we have

$$\rho_{g(T)}(\mathbf{X}_t) \le \mathrm{e}^{C_0 T} \left(\rho_{g(\tau_1 t)}(o, X_{\tau_1 t}) + \rho_{g(\tau_2 t)}(o, Y_{\tau_2 t}) \right).$$

Thus the proof can be reduced to the following claim:

$$\mathbb{E}\left[\sup_{\bar{\tau}_2 s \le t \le T} \rho_{g(\bar{\tau}_2 t)}(o, Y_{\bar{\tau}_2 t})^4 \right] < \infty. \tag{35}$$

Indeed, a similar bound for $X_{\bar{\tau}_1 t}$ follows in the same way. As shown in [12], $(\rho_{g(t)}(o, Y_t))_{[\bar{\tau}_2 s, T]}$ is dominated from above by a Bessel process (of dimension d) plus a constant. Thus (35) easily follows from the Burkholder inequality for the fourth moment of a Euclidean Brownian motion. □

Proof of Theorem 2. First we remark that the map $(x, y) \mapsto (X^\alpha, Y^\alpha)$ is obviously measurable. Thus, we obtain the same measurability for (X, Y). The integrability of $\Theta(t, \mathbf{X}_t)$ follows from Lemma 6. We will show the supermartingale property in the sequel. For $s \le s_1 < \cdots < s_m < t' < t < T$ and $f_1, \ldots, f_m \in C_c(M \times M \to \mathbb{R})$ with $0 \le f_j \le 1$, Set $F(\mathbf{w}) := \prod_{j=1}^m f_j(\mathbf{w}_{s_j})$ for $\mathbf{w} \in C([s, T/\bar{\tau}_2] \to M \times M)$. Take $\eta > 0$ arbitrarily and choose a relatively compact open set $M_0 \subset M$ so that $\mathbb{P}[\sigma^0_{M_0} \le t] \le \eta$ holds. Note that $\limsup_{\varepsilon \to 0} \mathbb{P}[\sigma^\varepsilon_{M_0} \le t] \le \eta$ also holds since $\{w \mid \sigma_{M_0}(w) \le t\}$ is closed. It suffices to show that there is a constant $C > 0$ which is independent of η and M_0 such that,

$$\mathbb{E}\left[\left(\Theta(t \wedge \sigma^0_{M_0}, \mathbf{X}_{t \wedge \sigma^0_{M_0}}) - \Theta(t' \wedge \sigma^0_{M_0}, \mathbf{X}_{t' \wedge \sigma^0_{M_0}}) \right) F(\mathbf{X}_{\cdot \wedge \sigma^0_{M_0}}) \right] \le C\sqrt{\eta} \tag{36}$$

holds. In fact, once we have shown (36), then Lemma 6 yields

$$\mathbb{E}\left[\left(\Theta(t, \mathbf{X}_t) - \Theta(s, \mathbf{X}_s) \right) F(\mathbf{X}) \right] \le 0$$

since $\sigma^0_{M_0} \to \infty$ almost surely as $M_0 \uparrow M$.

Take $f \in C_c(M \times M)$ such that $0 \le f \le 1$ and $f|_U \equiv 1$, where $U \subset M \times M$ is a open set containing $\bar{M}_0 \times \bar{M}_0$. Then, by virtue of Lemma 6 and the choice of M_0,

$$\begin{aligned}\mathbb{E}\Big[\Big(\Theta(t \wedge \sigma^0_{M_0}, \mathbf{X}_{t\wedge\sigma^0_{M_0}}) - \Theta(t' \wedge \sigma^0_{M_0}, \mathbf{X}_{t'\wedge\sigma^0_{M_0}})\Big)F(\mathbf{X}_{\cdot\wedge\sigma^0_{M_0}})\Big]&\\ \le \mathbb{E}\left[\left(\Theta(t,\mathbf{X}_t) - \Theta(t',\mathbf{X}_{t'})\right) f(\mathbf{X}_t) f(\mathbf{X}_{t'}) F(\mathbf{X})\ ;\ \sigma^0_{M_0} > t\right] + 2C_3^{1/2}\sqrt{\eta}.&\quad (37)\end{aligned}$$

For $u \in [s, T/\bar{\tau}_2]$, let us define $\lfloor u \rfloor_\varepsilon$ by

$$\lfloor u \rfloor_\varepsilon := \sup\{s + \varepsilon^2 n \mid n \in \mathbb{N} \cup \{0\}, 1 + \varepsilon^2 n < u\}.$$

Then, since $\{w \mid \sigma_{M_0}(w) > t\}$ is open,

$$\begin{aligned}&\mathbb{E}\left[\left(\Theta(t,\mathbf{X}_t) - \Theta(t',\mathbf{X}_{t'})\right) f(\mathbf{X}_t) f(\mathbf{X}_{t'}) F(\mathbf{X})\ ;\ \sigma^0_{M_0} > t\right]\\ &\le \liminf_{\varepsilon\to0} \mathbb{E}\left[\left(\Theta(t,\mathbf{X}^\varepsilon_t) - \Theta(t',\mathbf{X}^\varepsilon_{t'})\right) f(\mathbf{X}^\varepsilon_t) f(\mathbf{X}^\varepsilon_{t'}) F(\mathbf{X}^\varepsilon)\ ;\ \sigma^\varepsilon_{M_0} > t\right]\\ &= \liminf_{\varepsilon\to0} \mathbb{E}\left[\left(\Theta(\lfloor t\rfloor_\varepsilon,\mathbf{X}^\varepsilon_{\lfloor t\rfloor_\varepsilon}) - \Theta(\lfloor t'\rfloor_\varepsilon,\mathbf{X}^\varepsilon_{\lfloor t'\rfloor_\varepsilon})\right) f(\mathbf{X}^\varepsilon_{\lfloor t\rfloor_\varepsilon}) f(\mathbf{X}^\varepsilon_{\lfloor t'\rfloor_\varepsilon}) F(\mathbf{X}^\varepsilon)\ ;\ \sigma^\varepsilon_{M_0} > t\right]. \quad (38)\end{aligned}$$

Here the last inequality follows from the continuity of Θ and f. Set $\hat{\sigma}^\varepsilon_{M_0} := \lfloor \sigma^\varepsilon_{M_0} \rfloor_\varepsilon + \varepsilon^2$. Note that $\{\hat{\sigma}^\varepsilon_{M_0} = t_n\} \in \mathcal{G}_n$ for all $n \in \mathbb{N} \cup \{0\}$. Then

$$\begin{aligned}&\mathbb{E}\left[\left(\Theta(\lfloor t\rfloor_\varepsilon,\mathbf{X}^\varepsilon_{\lfloor t\rfloor_\varepsilon}) - \Theta(\lfloor t'\rfloor_\varepsilon,\mathbf{X}^\varepsilon_{\lfloor t'\rfloor_\varepsilon})\right) f(\mathbf{X}^\varepsilon_{\lfloor t\rfloor_\varepsilon}) f(\mathbf{X}^\varepsilon_{\lfloor t'\rfloor_\varepsilon}) F(\mathbf{X}^\varepsilon)\ ;\ \sigma^\varepsilon_{M_0} > t\right]\\ &\quad\le \mathbb{E}\left[\left(\Theta(\lfloor t\rfloor_\varepsilon \wedge \hat{\sigma}^\varepsilon_{M_0},\mathbf{X}^\varepsilon_{\lfloor t\rfloor_\varepsilon\wedge\hat{\sigma}^\varepsilon_{M_0}}) - \Theta(\lfloor t'\rfloor_\varepsilon \wedge \hat{\sigma}^\varepsilon_{M_0},\mathbf{X}^\varepsilon_{\lfloor t'\rfloor_\varepsilon\wedge\hat{\sigma}^\varepsilon_{M_0}})\right) F(\mathbf{X}^\varepsilon_{\cdot\wedge\hat{\sigma}^\varepsilon_{M_0}})\right]\\ &\qquad + 2\mathbb{E}\left[\sup_{s\le u\le T/\bar{\tau}_2} |\Theta(u,\mathbf{X}^\varepsilon_u) f(\mathbf{X}^\varepsilon_u)|^2\right]^{1/2} \mathbb{P}[\sigma^\varepsilon_{M_0} \le t]^{1/2}. \quad (39)\end{aligned}$$

Since a function $\mathbf{w} \mapsto \sup_{1\le u\le T/\bar{\tau}_2} |\Theta(u,\mathbf{w}_u) f(\mathbf{w}_u)|$ on $C([s,T/\bar{\tau}_2] \to M \times M)$ is bounded and continuous, we have

$$\limsup_{\varepsilon\to0} \mathbb{E}\left[\sup_{s\le u\le T/\bar{\tau}_2} |\Theta(u,\mathbf{X}^\varepsilon_u) f(\mathbf{X}^\varepsilon_u)|^2\right]^{1/2} \mathbb{P}[\sigma^\varepsilon_{M_0} \le t]^{1/2} \le C_3^{1/2}\sqrt{\eta}. \quad (40)$$

By combining (38), (39) and (40) with (37), the proof of (36) is reduced to show the following estimate:

$$\limsup_{\varepsilon\to0} \mathbb{E}\left[\left(\Theta(\lfloor t\rfloor_\varepsilon \wedge \hat{\sigma}^\varepsilon_{M_0},\mathbf{X}^\varepsilon_{\lfloor t\rfloor_\varepsilon\wedge\hat{\sigma}^\varepsilon_{M_0}}) - \Theta(\lfloor t'\rfloor_\varepsilon \wedge \hat{\sigma}^\varepsilon_{M_0},\mathbf{X}^\varepsilon_{\lfloor t'\rfloor_\varepsilon\wedge\hat{\sigma}^\varepsilon_{M_0}})\right) F(\mathbf{X}^\varepsilon_{\cdot\wedge\hat{\sigma}^\varepsilon_{M_0}})\right] \le C\sqrt{\eta}. \quad (41)$$

Take N_1 and N_2 so that $t_{N_1} = \lfloor t'\rfloor_\varepsilon \wedge \hat{\sigma}^\varepsilon_{M_0}$ and $t_{N_2} = \lfloor t\rfloor_\varepsilon \wedge \hat{\sigma}^\varepsilon_{M_0}$ hold. Let E_η be an event defined by

$$E_\eta := \left\{ \left| \sum_{n=N_1}^{N_2} \sqrt{t_n}(\Sigma_n - \bar{\Sigma}_n) \right| \le \frac{\sqrt{\eta}}{2\varepsilon^2(\sqrt{\bar{\tau}_2} - \sqrt{\bar{\tau}_1})} \right\}.$$

On E_η, an iteration of (32) together with (33) yields

$$\Theta(t_{N_2},\mathbf{X}_{t_{N_2}}) - \Theta(t_{N_1},\mathbf{X}_{t_{N_1}}) \le \varepsilon \sum_{n=N_1+1}^{N_2} \left(\sqrt{\bar{\tau}_2 t_n} - \sqrt{\bar{\tau}_1 t_n}\right)\zeta_n + 2\sqrt{\eta}$$

for sufficiently small ε. In addition, Lemma 5 yields $\limsup_{\varepsilon\to 0}\mathbb{P}[E_\eta^c] = 0$. By applying Lemma 4 with $T_0 = \bar{\tau}_2 t$ to an iteration of (32), we obtain a constant $C > 0$ satisfying

$$\left| \Theta(t_{N_2}, \mathbf{X}^\varepsilon_{t_{N_2}}) - \Theta(t_{N_1}, \mathbf{X}^\varepsilon_{t_{N_1}}) - \varepsilon \sum_{n=N_1+1}^{N_2} \left(\sqrt{\bar{\tau}_2 t_n} - \sqrt{\bar{\tau}_1 t_n} \right) \zeta_n \right| < C$$

uniformly in sufficiently small $\varepsilon > 0$. Since $F(\mathbf{X}^\varepsilon_{\cdot\wedge\hat{\sigma}_{M_0}})$ is $\mathcal{G}_{N_1}$-measurable, we obtain

$$\mathbb{E}\left[(\Theta(t_{N_2}, \mathbf{X}^\varepsilon_{t_{N_2}}) - \Theta(t_{N_1}, \mathbf{X}^\varepsilon_{t_{N_1}})) F(\mathbf{X}^\varepsilon_{\cdot\wedge\hat{\sigma}^\varepsilon_{M_0}}) \right] \le C\,\mathbb{P}\left[E_\eta^c\right]^{1/2} + 2\sqrt{\eta}.$$

Hence (41) holds with $C = 1/2$ and the proof is completed. □

References

[1] M. Arnaudon, K. A. Coulibaly, A. Thalmaier, Brownian motion with respect to a metric depending on time; definition, existence and application to Ricci flow, *C. R. Acad. Sci. Paris, Ser. I* **346** (2008), 773–778.

[2] M. Arnaudon, K. A. Coulibaly, A. Thalmaier, Horizontal diffusion in C^1 path space, arXiv:0904.2762, to appear in *Sém. Prob.*

[3] H.-D. Cao, X.-P. Zhu, A complete proof of the Poincaré and geometrization conjectures—application of the Hamilton-Perelman theory of the Ricci flow. *Asian J. Math.* **10** (2006), 165–492.

[4] B.-L. Chen, X.-P. Zhu, Uniqueness of the Ricci flow on complete noncompact manifolds, *J. Diff. Geom.* **74** (2006), 119–154.

[5] B. Chow, S.-C. Chu, D. Glickenstein, C. Guenther, J. Isenberg, T. Ivey, D. Knopf, P. Lu, F. Luo, L. Ni, *The Ricci Flow: Techniques and Applications, Part I: Geometric Aspects*, American Mathematical Society, Providence, Rhode Island, 2007.

[6] B. Chow, D. Knopf, *The Ricci flow: An introduction*, Mathematical Surveys and Monographs, vol.110, Amer. Math. Soc., Providence, Rhode Island, 2004.

[7] K. A. Coulibaly, Brownian motion with respect to time-changing Riemannian metrics, applications to Ricci flow, Preprint (2009), arXiv:0901.1999.

[8] R. S. Hamilton, Three-manifolds with positive Ricci curvature, *J. Differ. Geom.* **17** (1982), 255–306.

[9] B. Kleiner, J. Lott, Notes on Perelman's papers, *Geom. Topol.* **12** (2008), 2587–2855.

[10] K. Kuwada, Coupling of the Brownian motion via discrete approximation under lower Ricci curvature bounds, *Probabilistic Approach to Geometry* (Kyoto 2008), 273–292, Adv. Stud. Pure Math. vol. 57, Math. Soc. Japan, Tokyo, 2010.

[11] K. Kuwada, Coupling by reflection via discrete approximation under a backward Ricci flow, Preprint (2010), arXiv:1007.0275v1.

[12] K. Kuwada, R. Philipowski, Non-explosion of diffusion processes on manifolds with time-dependent metric, to appear in *Math. Z.*

[13] J. M. Lee, *Riemannian manifolds: An introduction to curvature*, Springer-Verlag, New York, 1997.

[14] J. Lott, Optimal transport and Perelman's reduced volume, *Calc. Var.* **36** (2009), 49–84.

[15] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, *Ann. Math.* **169** (2009), 903–991.

[16] R. J. McCann, P. Topping, Ricci flow, entropy and optimal transportation, *Amer. J. Math.* **132** (2010), 711–730.

[17] J. Morgan, G. Tian, *Ricci Flow and the Poincaré Conjecture*, American Mathematical Society, Providence, RI, 2007.

[18] G. Perelman, The entropy formula for the Ricci flow and its geometric applications, Preprint (2002), arXiv:math/0211159v1.

[19] G. Perelman, Ricci flow with surgery on three-manifolds, Preprint (2003), arXiv:math/0303109v1.

[20] G. Perelman, Finite extinction time for the solutions to the Ricci flow on certain three-manifolds, Preprint (2003), arXiv:math/0307245v1.

[21] W.-X. Shi, Deforming the metric on complete Riemannian manifolds, *J. Diff. Geom.* **30** (1989), 223–301.

[22] P. Topping, *Lectures on the Ricci flow*, Cambridge University Press, 2006.

[23] P. Topping, $\mathcal{L}$-optimal transportation for Ricci flow, *J. reine angew. Math.* **636** (2009), 93–122.

[24] C. Villani, *Optimal Transport, Old and New*, Grundlehren der Mathematischen Wissenschaften 338, Springer-Verlag, 2009.

[25] R. G. Ye, On the l-function and the reduced volume of Perelman I, *Trans. Amer. Math. Soc.* **360** (2008), no.1, 507–531.